Ecological Factors and Historical Biogeography Influence the Evolutionary Divergence of Insular Rodents

by

Paul A.P. Durst

Department of Biology
Duke University

Date:_____

Approved:

_____
V. Louise Roth, Supervisor

_____
Clifford Cunningham

_____
John Mercer

_____
A. Jonathan Shaw

_____
Paul Manos

Dissertation submitted in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy in the Department of
Biology in the Graduate School
of Duke University

2014

ABSTRACT

Ecological Factors and Historical Biogeography Influence the Evolutionary

Divergence of Insular Rodents

by

Paul A.P. Durst

Department of Biology
Duke University

Date:_____

Approved:

_____
V. Louise Roth, Supervisor

_____
Clifford Cunningham

_____
John Mercer

_____
A. Jonathan Shaw

_____
Paul Manos

An abstract of a dissertation submitted in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy in the Department of
Biology in the Graduate School
of Duke University

2014

Abstract

Islands have been the inspiration for some of evolutionary biology's most important advances. This is largely due to the unique properties of islands that promote the differentiation of island species from their mainland counterparts. Rodents are widely distributed across even the most remote islands, a rarity among mammals, making them uniquely suited to study the factors leading to the divergence of insular species. In this dissertation, I use two case studies to examine the morphological and genetic divergences that take place in an insular environment.

In chapters one and two, I examine how different factors influence insular body size change in rodents. In chapter one, I examine factors influencing the direction of island body size change using classification tree and random forest (CART) analyses. I observe strong consistency in the direction of size change within islands and within species, but little consistency at broader taxonomic scales. Including island and species traits in the CART analyses, I find mainland body mass to be the most important factor influencing size change. Other variables are significant, though their roles seem to be context-dependent.

In chapter two, I use the distributions of mainland rodent population body sizes to identify 'extreme' insular rodent populations and compare traits associated with those populations and their islands with those island populations of a more typical size.

I find that althought there is no trend among all insular rodents towards a larger or smaller size, 'extreme' populations are more likely to increase in size. Using CART methods, I develop a predictive model for insular size change that identifies resource limitations as the main driver when insular rodent populations become 'extremely small'.

Chapters three and four shift their focus to a single rodent species, the deer mouse *Peromyscus maniculatus*, as they examine the genetic differentiation of deer mice across the California Channel Islands and the nearby mainland. In chapter three, I sequence a region of the mitochondrial control region for individuals from 8 populations across the northern Channel Islands and two mainland sites, and I analyze these sequences by calculating population genetics parameters and creating a Bayesian inference tree and a statistical parsimony haplotype network. All of these analyses reveal significant divergences between island and mainland populations. Among the islands, Santa Barbara and Anacapa islands both display unique genetic signatures, but the other northern islands remain relatively undifferentiated.

In chapter four, I genotype individuals from the previous chapter at 5 microsatellite loci, I calculate additional population genetics parameters and I utilize a Bayesian clustering algorithm to examine the similarities and differences between nuclear and mitochondrial analyses. I find the nuclear data to be largely congruent with the mitochondrial analyses; there are significant differences between island and

mainland populations, and Anacapa Island is significantly differentiated from the other islands. Unlike the previous analyses, Santa Barbara Island is not significantly different from the northern islands, yet San Miguel Island has a unique genetic signature.

These studies underscore the importance of ecological processes and historical biogeography in the generation of diversity, and they highlight the role of islands as drivers of evolutionary divergence.

Dedication

To my parents, Mark Durst and Rachel Pinette, who have happily encouraged and supported all of my insane endeavors.

Contents

Abstract iv

List of Tables xi

List of Figures xii

Acknowledgements xiii

Introduction 1

1. Classification tree methods provide a multifactorial approach to predicting insular body size evolution in rodents 5

1.1 Introduction 5

1.2 Methods 6

1.2.1 Data Collection 6

1.2.2 Assessing the Phylogenetic and Within-Island Consistency of Size Change 8

1.2.3 Classification-Tree Analyses 11

1.3 Results and Discussion 14

1.3.1 Consistency of Size Change within Islands and Across Phylogeny 14

1.3.2 Classification-Tree Analyses 15

2. Mainland size variation informs predictive models of exceptional insular body size change in rodents 24

2.1 Introduction 24

2.2 Methods 27

2.2.1 Island Data Collection 27

2.2.2 Mainland Population Collection 28

2.2.3 Defining 'Meaningful' Size Change 28

2.2.4 Comparing Conditions associated with 'normal' vs 'extreme' Island Populations 29

2.2.5 Classification Tree Analyses 30

2.3 Results 31

2.3.1 Defining 'Meaningful' Size Change 31

2.3.2 Comparing Normal and Extreme Islands 32

2.3.3 Classification Tree Analyses 33

2.4 Discussion 35

2.5 Reflections on Chapters One and Two 40

3. A mitochondrial phylogeography of *Peromyscus maniculatus* across the northern California Channel Islands 43

3.1 Introduction 43

3.2 Methods 47

3.2.1 Sampling 47

3.2.2 Molecular Genetics 47

3.2.3 Population Genetic, Phylogeographic and Phylogenetic Analyses 48

3.3 Results 50

3.4 Discussion 57

4. A comparison between mitochondrial sequence data and microsatellite loci for *Peromyscus maniculatus* on the California Channel Islands 61

4.1 Introduction 61

4.2 Methods 63

4.2.1 Sampling and Molecular Genetics 63

4.2.2 Microsatellite Analysis 63

4.3 Results 64

4.3.1 Genetic Diversity Analyses 64

4.3.2 Genetic Structure Analysis 65

4.4 Discussion 69

4.5 Comparisons with other Channel Islands species 70

4.6 Future Directions and Conclusions 77

Appendix A. Reading the classification tree in Chapter One. 79

Appendix B. Supplementary tables and figures for Chapter One. 80

Appendix C. Assessing the phylogenetic structure of size change in Chapter Two 83

Appendix D. Examining the asymmetric boundary for 'Extreme' size change in Chapter Two 86

Appendix E. Supplementary table for Chapter Two 89

References 90

Biography 103

List of Tables

Table 1. Pairwise F_{ST} estimates from the mitochondrial control region for populations in this study. 54

Table 2. Differentiation of the populations in this study as determined by an analysis of molecular variance (AMOVA). 55

Table 3. The number of alleles (A), allelic richness (R_S), number of private alles (priv), expected heterozygosity (H_e) and observed heterozygosity (H_O) for each population in this study. 66

Table 4. Pairwise F_{ST} estimates from 5 microsatellite loci for populations in this study. 67

Table 5. Loadings of original BioClim variables on PC1 of principal component analyses (using correlation matrix) for precipitation and temperature. 80

Table 6. Effects of Shifting the "Extreme" Boundaries 88

Table 7. Loadings of island terrain heterogeneity variables on PC1 of principal component analysis 89

List of Figures

Figure 1. Cross-validation plot for the classification tree. 13

Figure 2. Relative importance (±SD) of predictor variables from the random forest analysis of size change. 16

Figure 3. Classification tree showing predicted direction of size change for insular rodents. 18

Figure 4. Comparing distributions of numbers of island populations with various characteristics. 33

Figure 5. Classification tree, showing predicted direction of size change for insular rodent populations, and body size ranges corresponding to its terminal nodes. 34

Figure 6. Timelines of important events in Channel Islands history. 44

Figure 7. Map of the California Channel Islands. 46

Figure 8. Site frequency spectrums. 50

Figure 9. Haplotype network constructed from mitochondrial control region sequence data for northern Channel Island Peromyscus maniculatus and its mainland relatives. 53

Figure 10. Bayesian phylogenetic tree for the haplotypes of mitochondrial control region sequences present in this study. 56

Figure 11. Inferred clusters based on a STRUCTURE analysis for K=5 clusters. 68

Figure 12. A phylogenetic tree of all rodent species used in this study. 81

Figure 13. Stability of classification tree using different climate variables. 82

Figure 14. Phylogenetic relationships among species examined in chapter two based on the phylogenetic tree depicted in Figure 12. 85

Acknowledgements

Coming in to graduate school, I knew that biology was supposed to be collaborative, but I had no idea the extent to which that would be the case as I moved toward my PhD. There are so many people I would like to thank and, while a few pages of acknowledgements would definitely add some much needed heft to this dissertation, there are only so many trees in the world. Since at least one copy of this document will be printed out, I'll try to keep it brief.

First and foremost, I would like to thank my advisor, Louise Roth. Her wisdom and guidance have been critical to my development as a biologist. I will be forever grateful to her for taking me under her wing. I would also like to thank my committee members past and present, John Mercer, Cliff Cunningham, Paul Manos, Jon Shaw and Bill Morris, for their challenging and thoughtful insight that has brought me to this point.

I would like to thank Shai Meiri, who made the island body size projects possible by graciously sharing his data with the world, the Natural History Museum of Los Angeles County (especially Jim Dines) and the Smithsonian, U.S. National Museum of Natural History for making their collection databases available, and Tim Coonan and the rangers and researchers at Channel Islands National Park, Paul Collins at the Santa Barbara Museum of Natural History, Lyndal Laughrin at the Santa Cruz Island Reserve,

Kate McCurdy at Sedgwick Reserve, Diana Roemer from Ventura's office of Parks, Recreation and Community Partnerships and my field technicians, Ben Hartman and Genevieve Durst, for their support in the field.

I would also like to thank members of the Duke Biology community who generously donated their time and efforts to the cause, listening to my ideas, explaining basic biology concepts to me and helping me troubleshoot when my lab work inevitably failed. I am grateful to the members of the NESCent Mammal Reading Group (especially Trina Roberts) and Island Reading Group for being the lab-mates I never had, the Duke Biology staff, especially Anne Lacey, Jo Bernhardt, Jim Tunney, Randy Smith and Jill Foster for making sure I had the academic and financial resources I needed, Matt Johnson for joining me in Friday afternoon procrastinations, Dudu Meireles for telling me why I was wrong, and Carrie Olson-Manning who, by yelling at me to do work, saved the Channel Islands project on several occasions.

Official funding for this project came from Duke Biology, a small grant in aid from Sigma Xi, a graduate fellowship from NESCent, National Science Foundation EF-0905606 and a National Science Foundation DDIG # 1311597. A special thanks to those who supported my SciFund crowd-sourced funding initiative, especially for the generous support of Jane Macfarlane, Dave and Pat Lea, Mark Durst, Rachel Pinette, Micheline LeBlanc, Rev. Peter McGrath, Elizabeth Durst, Annette Counts, Genevieve Arechiga and Marian Sandoval along with the support of the Louise Roth discretionary

fund. I would also like to thank the Mark Durst and Rachel Pinette Science Foundation, whose contributions ensured that my statistics were rigorous, and my summers fruitful.

Finally, I am especially grateful to Amanda Lea, whose encouragement, assistance and support made this dissertation possible, and whose love, care and companionship have made the last two years wonderful.

Introduction

For over 100 years, islands have been a source of fascination and inspiration for scientists like few other places on earth. Islands have been the drivers of many important advances in evolution and ecology from Darwin and Wallace's theories of natural selection and biogeography (Darwin 1859; Wallace 1876) to more recent advances in community ecology, and conservation biology (Adler and Levins 1994; Wang, Chen, and Ding 2011; Boyer 2010; Pergams, Lacy, and Ashley 2000). One reason islands have had such undue influence in biology over the last century is because islands act as natural laboratories for evolution and ecology. In these fields, where large scale manipulations are nearly impossible, islands provide a system free of many of the confounding factors found on the mainland, allowing scientists to develop new theories, and test them across a wide range of biotic and abiotic conditions (Whittaker and Fernández-Palacios 2007).

One of the most noteworthy ways islands inform evolutionary biology is in their role as generators of diversity. As natural laboratories, islands have produced some of the best-studied examples of evolutionary divergence and differentiation such as Darwin's finches, Galapagos tortoises, and Carribean anoles. While islands are generally species poor, they have much higher rates of endemism than seen on the mainland (Barraclough, Vogler, and Harvey 1998). In fact, of the 35 regions designated as biodiversity hotspots by the Critical Ecosystem Partnership Fund (one of the criteria

being high rates of endemism, the other being habitat destruction), more than a third of them are made up, at least in part, of island ecosystems (*CEPF Report: A New Model for Global Conservation* 2011). Depauperate species pools, smaller effective population sizes and increased isolation all contribute to the phenotypic and genetic changes that lead to high rates of endemism (Walker 1980). These conditions make it possible to study diversity-generating processes such as niche shifts, character displacement and speciation in a way that would otherwise be very difficult in mainland environments.

Among mammals, rodents provide a unique opportunity to study evolutionary divergences on islands. Excluding bats, rodents make up the majority of native insular mammals (there are over 70 species of rodents found on the islands of the Philippines alone; Heaney et al. 2002), and invasive rodents have spread to some of the most remote islands as stowaways on human vessels (Matisoo-Smith et al. 1998). Recent studies have demonstrated that insular rodents can undergo rapid change both in terms of phenotypic (Pergams and Ashley 1999; Millien 2006; Millien 2011) and genetic differentiation (Britton-Davidian et al. 2000). Their potential for rapid differentiation along with the continued use of rodents as model organisms in a mainland setting (Weber et al. 2010) means that insular rodents are uniquely positioned to address some of the long-standing questions about the factors contributing to phenotypic and genetic divergences.

In this thesis, I use insular rodents to explore two very different case studies examining islands as drivers of evolutionary divergence. The first case study is a synthetic analysis of insular body size in rodents. Synthetic studies are increasingly important for studying evolutionary processes because they take advantage of far more data than can be produced in a single laboratory- or field-based study (Sidlauskas et al. 2010). Integrating large amounts of data from a variety of sources makes it easier to examine large-scale questions, and can reveal patterns that would not be evident from a single more focused study. In chapters 1 and 2 of my dissertation, I take this broad approach, synthesizing hundreds of records of island rodent populations in an attempt to better understand the insular morphological changes known as the island rule. In Chapter 1, I use classification trees and random forests to examine the factors influencing the direction of insular size change. In Chapter 2 I focus on insular rodents that have undergone "extreme" size change, looking at factors that lead to extreme size change and creating a model to predict the directionality of changes in extreme size.

The second case study focuses in on a single species in one location, examining the genetic differentiation of the deer mouse, *Peromyscus maniculatus*, across the Channel Islands in southern California. The Channel Islands provide an excellent background for the study of evolutionary divergences. The islands occupy a unique position as recent, temperate continental islands that have never been connected to the mainland (Garth 1965) and a wealth of information exists regarding the geology, biology and

archaeology of this system (including 8 symposia spanning fifty years' worth of research). By integrating bathymetry data, archaeological records and molecular data, previous researchers have produced hypotheses about the patterns of evolutionary divergence we would expect to see across the islands (Gill 1976; Collins 1982; Ashley and Wills 1987). In Chapter 3, I test these hypotheses for *Peromyscus maniculatus* by constructing a molecular phylogeography of Channel Islands deer mice using mitochondrial sequences. In Chapter 4, I build on that phylogeography with a preliminary analysis of nuclear microsatellite loci in the same system.

Chapter 1 has been published in a peer-reviewed journal, and Chapter 2 is currently in the process of being submitted. Both of these chapters were written in collaboration with Louise Roth. I was responsible for the initial study design of the chapters, along with the majority of the computational analyses and manuscript writing, but as the studies expanded beyond their initial scope, Louise's input on further analyses and her contributions to sections of the manuscript were essential.

1. Classification tree methods provide a multifactorial approach to predicting insular body size evolution in rodents

Durst, P.A.P. & Roth, V.L. 2012. Classification tree methods provide a multifactorial approach to predicting insular body size evolution in rodents. *American Naturalist*.

1.1 Introduction

The "island rule" for mammalian body-size evolution—a pattern of size increases and decreases dramatically exemplified in the fossil record of Mediterranean islands by miniature elephants and giant dormice—is known for both its regularity and its exceptions. For instance, while dwarfing of large ungulates has occurred in parallel on multiple islands with law-like regularity, patterns of size change for rodents are less clear (Meiri, Cooper, and Purvis 2008). Proposed predictors of the direction of these size changes have included the species' original body size (Mark V Lomolino 1985) or phylogenetic affinities (Foster 1964; Meiri, Cooper, and Purvis 2008); island area (Heaney 1978), the number of predators or competitors (Raia and Meiri 2006), and other attributes of the habitat (Lawlor 1982). With no single explanation capable of predicting all instances of insular body-size change, Meiri et al. (2008) concluded that no rule exists, and that body-size evolution on islands is a function of historical and ecological contingency.

The variables that have been used to explain insular size change are characteristics of the species and of the island, but to what extent does the direction of size change depend on which species and which islands are involved? Additionally, to what degree is the dataset structured by the phylogenetic relationships among the taxa within it? Traditional linear, correlative methods can be difficult to apply to combinations of continuous and categorical variables, especially if the roles of the variables are context-dependent. Moreover, if a process is multifactorial, its causes may not be revealed by correlations.

The order Rodentia has presented particular challenges to the application of any 'rule' because among rodents on islands exist cases of both larger and smaller forms, and previous attempts to disentangle potential explanations have not dealt with many factors simultaneously. In this paper we first ask whether the direction of size change in insular rodents is consistent within islands and within taxa. We then make use of classification trees to identify which among an array of 'predictor' variables are most useful in dichotomously categorizing populations of island rodents that in comparison to their mainland relatives are "large" or "small".

1.2 Methods

1.2.1 Data Collection

We assembled 135 records (a record or case that is treated here as a unit to be classified is the population of a particular species on a particular island) of insular

rodents (73 species on 55 islands from across the world) building on a collection by Meiri et al. (2008) that included island:mainland body-size ratios, island areas, number of carnivoran species found on the island, and distances to the nearest mainland. To this information we added for each population (a) its taxonomic family according to Wilson and Reeder (2005), (b) numbers of con-familial species (assumed to be its competitors) on the same island, (c) diet information, (d) mainland masses according to Smith et al. (2003); and for each island (e) numbers of rodent species, and (f) 19 bioclimatic variables from the WorldClim database (Hijmans et al. 2005), which we summarized as PC1 from each of two principal component analyses (on temperature and on precipitation variables; see Appendix B).

Most of the dietary information used in this study was collected by Samantha Hopkins and Louise Roth as part of a larger database compiled for another study (Price et al. 2012) for all orders of mammals. On the basis of uniform criteria applied to records compiled from the primary literature, species were classified into trophic categories as carnivore, granivore, herbivore, or omnivore.

To determine the number of rodent species on each island, we found literature sources by searching Google Scholar with the terms "mammal checklist" and "mammal list" and the name of each island. The largest number of rodent species listed in these publications was used as the number of rodent species for the island, and the number of confamilials for each species/island record in our data set was tabulated from this list. In

cases where no list of mammals could be found for an island, we used the number of rodent species on the island that were measured by Meiri et al (2008).

We used the categorical variable "Family" as a rough qualitative assessment of phylogenetic relationship, natural history, and overall phenotypic or ecological similarity. Use of this variable was not intended, in itself, to provide information about interspecific phylogenetic structure in the dataset. This will be considered explicitly in the following section.

1.2.2 Assessing the Phylogenetic and Within-Island Consistency of Size Change

To examine phylogenetic structuring of the tendency to increase or decrease in size on islands, we carried out a permutation test. For all species in the dataset we assembled a phylogenetic tree (Figure 12) with a topology obtained by splicing together trees from multiple sources (Conroy and Cook 2000; Oshida 2004; Mercer and Roth 2003; Song et al. 2012; Jaarola et al. 2004; Jansa and Weksler 2004; Steppan, Adkins, and Anderson 2004; Alexander and Riddle 2005; Bradley et al. 2007; J. Michaux, Chevret, and Renaud 2007; Blanga-Kanfi et al. 2009; Fritz, Bininda-Emonds, and Purvis 2009). Each species was represented by a single terminal branch, and the character state of each terminal taxon (all populations increase in body size, all populations decrease in body size, mixed) was mapped onto the tree. Parsimony was used (in Mesquite; Maddison and Maddison 2011) to reconstruct ancestral states and to obtain the total length of the tree in character-transition steps. The distribution of terminal states across the tree was

then randomly permuted 1000 times, retaining the respective frequencies of the three character-states among the terminal taxa. Ancestral states and total tree-length were assessed for each permutation to produce a distribution of tree-lengths. The length of the original tree was compared to the distribution. Phylogenetic structuring of size changes would be indicated if the tree required a significantly smaller number of steps than expected for trees generated at random.

Phylogeny is relevant to this study at both inter- and intraspecific levels. At the intraspecific level, 24 species were represented in our dataset by more than one island population—a subset of cases we refer to as the "replicated populations". These replicated populations allowed us to assess whether distinct populations within a single species responded to their island environments with size change in the same direction. In nine species the direction of all changes, either dwarfing or gigantism, was the same for all populations, but the response of fifteen species was mixed, to varying extents (i.e., in some populations size increased and in others it decreased).

The tendency for a single pattern to predominate within a species was assessed by calculating the probability of finding clustering within species to the degree observed in the dataset (or higher), if the known increases and decreases were distributed among all the replicated populations at random. The probability was estimated in the following way: A tree diagram was constructed consisting of a basal polytomy of 24 branches (representing the species); each of these branches bore an additional, more distal

polytomy whose number of branches corresponded to the number of island populations in one of the replicated species, for a total of 90 terminal taxa. (This procedure was intended to substitute for an exact calculation of binomial probabilities: the basal polytomy represented independence among species; the terminal polytomies represented independent events of size change among populations.) The character-state of the binary variable "direction of size change on the island" (i.e., larger or smaller) was identified for each population, parsimony was used to reconstruct an ancestral state for each species and for the basal node, and the total length of the tree in state-change steps was obtained. The terminal states were then randomly permuted 1000 times on this topology and the tree-length recorded for each permutation. The length of the original tree diagram was compared to the distribution of lengths generated by permutation; a relatively short tree would indicate that size change is highly consistent within species.

To examine whether the observed directions of insular body size change were more consistent among species on a given island than would be predicted by chance alone, an analogous permutation test was carried out among islands for which more than one rodent species was recorded. Here, the basal polytomy consisted of 36 branches, representing each island with multiple species in the dataset, and for each island the number of rodent populations was represented in a polytomy of terminal branches. Again, the length of the original tree diagram was compared to the

distribution of lengths generated by 1000 permutations, and here a relatively short tree would indicate that size change is highly consistent within islands.

1.2.3 Classification-Tree Analyses

We used decision trees to predict whether species would increase or decrease in size on a given island. Decision-tree models (specifically classification and regression trees), which have been applied effectively to questions in ecology and conservation biology (Jones, Fielding, and Sullivan 2006; Davidson et al. 2009), are used to identify predictor variables that correctly classify samples according to a response variable. The main difference between classification and regression trees is that classification trees predict membership among the different categories of a qualitative variable whereas regression trees predict membership along a continuous variable.

With different assumptions and response variables, either method can be used to examine insular body size change. For example, regression trees could be used with island:mainland body-size ratios if it is assumed that meaningful distinctions can be inferred from different degrees of change. Yet unless the time of colonization or isolation is known for all cases and included in the model, use of a size ratio (or other measure of degree) assumes that differences in the value of the ratio reflect important differences in the process, rather than, for example, different stages in a time course of the process of size change. If size change on an island takes place over a period of time, differences in the degree of change may simply be due to the fact that more time has

passed for one island group—or even to subtle differences in the choice of the mainland population for comparison (Millien and Damuth 2004). Alternatively, classification trees can be used with direction of change—larger or smaller—as the response variable. While disregarding magnitude, this approach assumes that the direction, regardless of degree, is meaningful and contains information about the mechanisms of change. The decision tree we present is a classification tree.

Using the *rpart* package in R (R Development Core Team 2013; Therneau, Atkinson, and Ripley 2006), we constructed a fully-grown classification tree (no nodes pruned, minimum number of samples required for a split = 2) where each unique sample had its own terminal node. To determine the optimum number of nodes for the tree (to minimize over-fitting), we examined the cross-validation plot (Figure 1). The point after the graph drops below the dotted line (where the tree's error rate is within 1 SE of the minimum error tree) was interpreted as representing the largest number of meaningful nodes. Using this value, we pruned the fully-grown tree (pruning all nodes with a complexity parameter < 0.043, minimum number of samples required for a split = 2) to obtain the optimal tree.

Because classification trees can be sensitive to small changes in the data (Davidson et al. 2009) we also used the randomForest package in R (Liaw and Wiener 2002) to produce a random forest (bootstrapped subsets of the data) of 10,000 classification trees. Random forest methods use random subsets of the data to generate a

set of trees that are then used to assess the influence of single predictor variables and to generate a predictive model for the dataset (Cutler et al. 2007). Most studies that have used predictive models from random forests have used datasets with several thousand samples (e.g. Davidson et al. 2009; Boyer 2010). Considering the size of our dataset (135 records), we used the random forest primarily to generate a statistical assessment of variable importance.

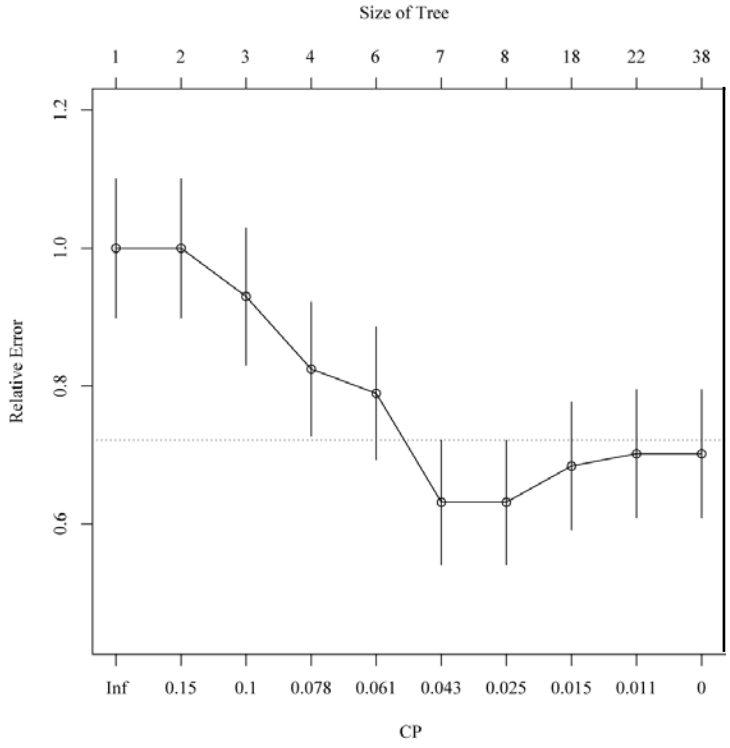


Figure 1. Cross-validation plot for the classification tree. The top y axis refers to the number of splits in the tree, and the lower y axis refers to the corresponding complexity parameter for the best tree of that size. The optimum number of nodes was determined to be the first value within one SE of the minimum error tree.

1.3 Results and Discussion

1.3.1 Consistency of Size Change within Islands and Across Phylogeny

The length of the original phylogenetic tree fell well within the distribution produced by random permutations ($P = 0.10$). Accordingly, we found little evidence of phylogenetic structuring of the interspecific distribution of size increases and decreases within the dataset examined here. (What structure there is appears to occur among the most closely related species: using polytomies in place of resolved topologies within genera in the tree yielded even lower significance, $P=0.30$)

For the intra-specific analysis, less than 2% of the random permutations produced trees shorter than the original (actual) groupings; i.e., $P < 0.02$, indicating that although the responses of replicate populations within a species were not uniform, they were consistent to a significant degree: the tendency was high for body sizes of all or most replicate island populations within a species to differ from relatives on the mainland in the same direction. We can therefore anticipate that organismal traits that are distinctive at the species level have an important influence on whether body size in a population is larger or smaller on islands.

The within-island analyses also revealed significant structuring ($P = 0.05$), suggesting that the characteristics of an island itself (potentially including both physical and biotic factors) influence in a relatively consistent way whether the populations of

rodents it supports are bigger or smaller in body size than their closest relatives on the mainland.

From these permutation tests it is clear that confronting a phenomenon like size change in insular rodents calls for a technique that makes use of a variety of types of predictive factors, including traits of both the species and the particular island.

1.3.2 Classification-Tree Analyses

The random forest model correctly classified 70.4% of the body-size changes (74.4% of size increases and 63.2% of size decreases, Cohen's kappa = 0.39, $P < 0.001$) while the classification tree correctly classified 83% of the changes (92.3% of the increases and 70.2% of the decreases, Cohen's kappa = 0.641, $P < 0.001$). The analyses were complementary: the random forest identified important variables associated with size change and the classification tree highlighted the interactions between these variables. In the random forest analysis, mainland body mass was the most important factor for predicting the direction of change, with most large rodents getting smaller and small ones tending to enlarge, but island area, climate, family, and the number of rodent species on the island were also significant (Figure 2). The island rule is typically framed in terms of a species' mainland mass so it may be no surprise that mass (and presumably its life-history correlates; Adler and Levins 1994) was found to be the most important predictor. This finding was reinforced by the classification tree, which used mass in its second split. Masses in our sample ranged from 6 g to 21,820 g (median=87.5 g,

mean=547 g), yet the mass used for this split (253 g) is remarkably close to the value of 272 g that Lomolino (2005) proposed as the size on which rodents converge,

(Nevertheless, see Raia, Carotenuto, and Meiri 2010 for arguments against optimal body sizes.) On its own, mass correctly predicts 65.2% of the size changes, approaching in its accuracy the entire random forest model.

Precipitation and temperature were important to the random forest analysis, and their importance was consistent regardless of whether they were expressed as many separate variables or in aggregate, as PC scores (see appendix C). Climatic variables may exert their influence on body size indirectly, through their effects on resource availability (Heaney 1978; Van Valen 1973), community structure, soil quality, or other factors. These factors, as well as other island-specific traits (e.g., topological complexity,

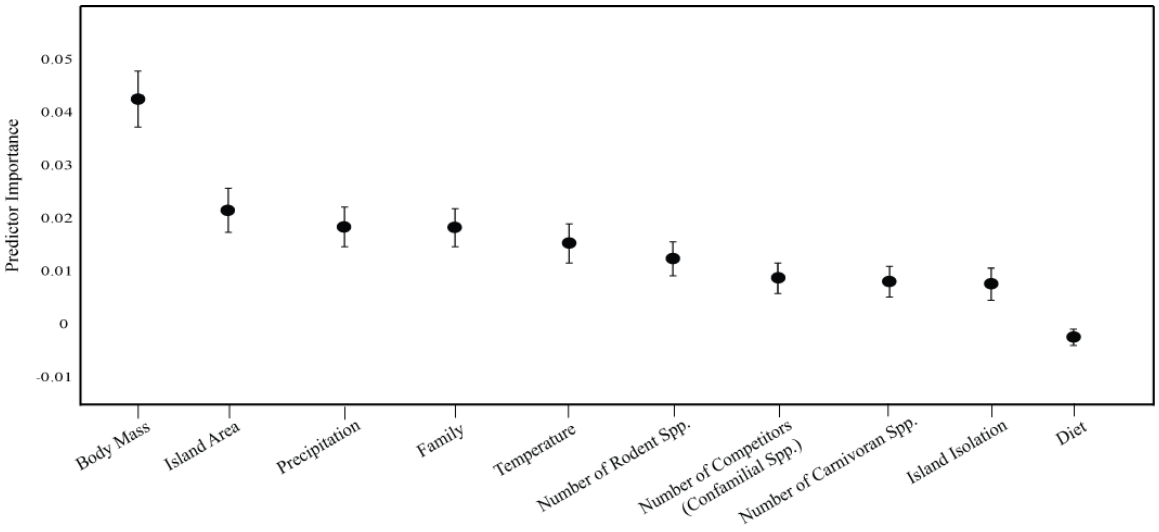


Figure 2. Relative importance (±SD) of predictor variables from the random forest analysis of size change. Importance is measured as the drop in accuracy from the removal of individual variables.

number of biomes, species richness, etc.) may be the sources of consistency in size change within islands that was identified here through permutation analysis. Each island provides a unique environment and individual islands cannot be viewed as interchangeable.

The classification-tree model both highlighted the interactions between explanatory variables and provided ranges of values where these interactions were important (Figure 3). For example, Lomolino (2005) suggested that as island area increases, islands become more similar to mainland habitats. Meiri et al. (2008) acknowledged this in compiling the original dataset, omitting records of species on islands larger than 50,000 km^2 (roughly, the area of Costa Rica), but this limit for island size was based on studies of carnivorans and large herbivores. An organism's ability to recognize an island should depend on the size of its home range, and mammalian home ranges are strongly affected by body size (Gaston and Blackburn 1996). The cutoff for what appears to be an island for rodents, which are generally smaller than carnivorans or artiodactyls, should be much smaller than 50,000 km^2. This was reflected in our classification tree, where small rodents (< 253 g) on islands smaller than 2,326 km^2 showed a strong trend toward size increase whereas island area on its own could not predict patterns of size change for small rodents on larger islands or for large rodents in general.

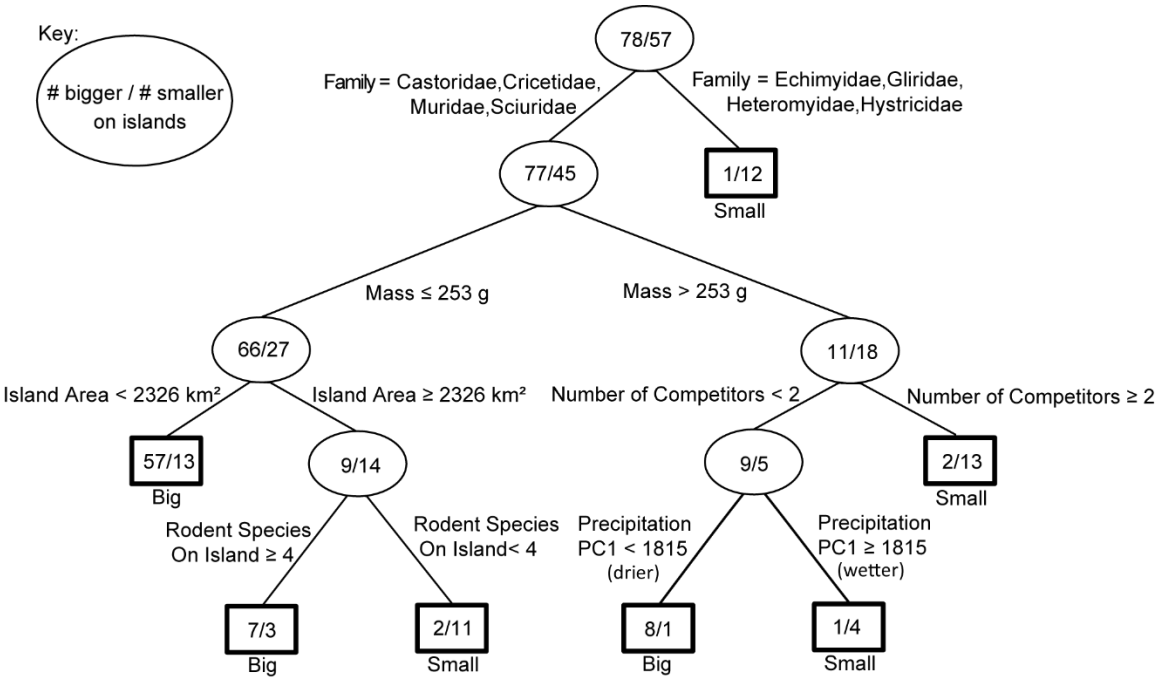


Figure 3. Classification tree showing predicted direction of size change for insular rodents. Numbers shown in each oval refer to the number of cases in the dataset (that are respectively bigger/or smaller on the islands) that fit the criteria described on the branches above that node. Rectangles indicate terminal nodes, with the predicted direction of change noted below. (For further description of how to "read" the tree see appendix A)

The classification tree also pointed to patterns within the data requiring additional research. Although the random forest identified mass as the most important predictor, and the permutation test indicated that phylogenetic structuring above the species level was weak (at best), family was the first split on the classification tree. Family can be considered a proxy for natural history, and despite our phylogenetic findings one might expect closely related families with similar natural histories to group together. However, for families that grouped together here this was not the case (Figure 3). Rather, the split along family lines is largely due to the heteromyid rodents, which

together constitute more than half the observations on the right (size-reduced) side of the split. Despite being well below the 253 g threshold, all heteromyids in our sample buck the trend predicted by mass and become smaller on islands. (Size reduction in heteromyids will be discussed further below in the context of diet.) Heteromyids' tendency to become smaller on islands is much more consistent than changes found in any other well-sampled family in our analysis (cf Lawlor 1982; Meiri et al. 2008): in fact, within only 4 of the 17 genera with more than one record in our dataset did all species change size in the same direction. Because of the consistency among heteromyids, the classification-tree analysis identified a split along family lines as the most meaningful split in the data, and several poorly sampled families that got smaller on islands were grouped with the heteromyids even though the biological explanation for this grouping is not clear. Removing family from the classification-tree analysis only caused other variables (specifically, mass and climate) to group all heteromyids together, joining them with a different set of small species that inhabit dry, seasonal islands; excluding heteromyids from the analyses eliminated the importance of family altogether.

Our results are consistent with several proposed hypotheses, but not every hypothesis received support. Lomolino (2005) suggested that release from certain ecological pressures was a major driver of size change on islands. For large herbivores, Raia and Meiri (2006) found that body size on Mediterranean islands could largely be explained by interspecific competition, although predation too was

important. Ecological interactions are difficult to measure directly, but several explanatory factors in our analyses were intended indirectly to account for them. Numbers of rodents, confamilial species, and carnivorans on the islands ranked low in relative importance (respectively 6th - 8th among 10 variables used in the random forest, Figure 2) but numbers of rodents and confamilial species proved useful to the classification tree analysis. Beyond the first split in the tree, the greatest number of cases successfully predicted to undergo size reduction are large species on islands with more than two confamilial (competitor) species, and small species on large islands with few other species of rodents. These patterns do not fit neatly with one ecological explanation that has accounted for size reduction in large insular ungulates: Raia, Barbera, and Conte (2003) interpreted such changes as a shift in life-history tactics along a "fast - slow" continuum (in large ungulates a shift to a faster lifestyle, emphasizing reproduction over somatic growth, accompanies size reduction, and is especially marked on islands that are low in species richness, especially of competitors, and experience relatively unstable and unpredictable conditions), but this framework is more difficult to apply to the patterns we observe in rodents. Reciprocally, and more specifically for insular rodents, Adler and Levins (1994) defined an "island syndrome" in which high and stable population densities produce size increase. This hypothesis was not proposed to account for size decreases, and further assessment of its role in insular size increases awaits more population-level data than are currently available. Arguments that emphasize

competitive displacement or ecological release are compromised by the fact that the largest and smallest rodent species on any given island in our analysis were just as likely to have converged in size (19 instances) as they were to diverge (17 instances). The role of interspecific interactions remains unclear and, to be characterized accurately, may require information on population density and consideration of the entire fauna, including other phyla (Brown and Davidson 1977).

Although some variables received little support from our analyses, they need not immediately be discounted. Lawlor (1982) attributed size reduction in insular heteromyids to their specialization on limited and coarse-grained food supplies (seeds), and the limited number of specialist species in our dataset showed very consistent patterns of size change: the three carnivorous rodents all increased in size while among the herbivores the three granivores all decreased in size. Yet diet ranked lowest among variables in the random forest analysis (Figure 2). It may be that dietary specialization influences insular body size change, but the relatively broad categorization of diet used here and the small number of true dietary specialists among the rodents (Landry 1970) make it difficult to detect a pattern. Dietary specialization can occur on a local scale and resource availability differs from island to island, but the diet data were collected to represent each species across its entire range. Consistency in the direction of size change for different populations of single species in our dataset was high, but not perfect. A species that is relatively flexible in its diet could be obliged to specialize on grains on one

island (for example) while including softer fruits on another, potentially leading to differences in the direction of change for the two populations. Population-level diet data are not available for the island populations in our sample, but they might reveal a more important role for diet than our analyses suggest.

Predation is another variable found to be unimportant that should not be fully discounted. While the number of carnivorans on an island did not seem to influence body size significantly, avian and reptilian predators, which were not considered here, may have a larger impact on insular rodents (Michaux et al. 2002). The number of carnivorans on an island may be more important for larger mammals (Raia and Meiri 2006) that are less likely to be preyed upon by birds or reptiles.

The island rule is a complex, intricate problem that has defied any simple explanation. Analyzing the multifactorial processes leading to size change on islands benefits from a method that is not constrained to detecting linear or monotonic relationships--one that combines categorical, qualitative, and quantitative factors and can consider their interactions in different contexts. In our analysis of rodents, permutation tests demonstrated that within-island and within-species processes each produce consistent patterns in the directions of change, but the lack of significant patterning at higher taxonomic levels suggested that within rodents these patterns are not otherwise phylogenetically constrained. By applying classification-tree methods to the same data, we found that body mass was the strongest predictor of whether size

increases or decreases: the classic phrasing of the island rule, that small things get big while big things get small, tends to hold true within rodents. It is on very small islands (<2326 km^2) that small rodents show the most consistent tendency to enlarge, and in contrast to earlier findings for large herbivores on Mediterrantean islands (Raia and Meiri 2006), ecological interactions, as they were represented in our analyses, were comparatively uninformative: The smallest and largest rodents on an island converged or diverged in size with equal frequency, the number of carnivoran species on an island was unimportant as a predictor, and a fauna depauperate in other species of rodent and/or members of the same family yields, if anything, a slight tendency to buck the trend predicted by size (*i.e.* overall, 9/17 cases of large rodents in this situation get larger, and 20/36 cases of small rodents get smaller).

Identifying other relevant factors to account for unexplained variation presents a challenge for future work. The approach illustrated here, we aver, provides a promising framework for assessing the importance of additional variables and incorporating them into the predictive model of insular body-size change.

2. Mainland size variation informs predictive models of exceptional insular body size change in rodents

Durst, P.A.P. & Roth, V.L. 2014. Mainland size variation informs predictive models of exceptional insular body size change in rodents. In revision.

2.1 Introduction

The "island rule" for mammals, a pattern of divergence in body size between insular populations of mammals and their mainland counterparts, was once deemed "an extraordinary phenomenon which seems to have fewer exceptions than any other ecotypic rule in animals" (Van Valen 1973). Yet as examples have proliferated—from anecdotal observations of insular mammals of unusual size (Busk 1868, Adams 1874), to tabulations and tallies of taxa showing insular gigantism or dwarfism (Foster 1964), to regressions on body mass of the continuous variable "size ratio" (average island / average mainland body mass of a species; (Mark V Lomolino 1985)—so too have the exceptions. Shifts in the average body size of a population of insular mammals have been observed to differ in direction and in degree for different populations on different islands. And in recognition of this, the number of factors proposed to have a causal role in the changes has also grown. The simple scenario that Foster (1964) described, in which size increase or decrease predominates within a given order of mammals, has been replaced with the description of a generally monotonic trend of decrease in size ratio with increasing body size across species (Mark V Lomolino 1985). This, in turn, has been refined and replaced by models of size change that apply differently in different

cases and are contingent both on the phylogenetic affinity of the population and on an array of ecological, environmental, geographic, and species-specific attributes of the island + population pair (Meiri, Cooper, and Purvis 2008; Durst and Roth 2012; McClain et al. 2013; Lomolino et al. 2012; Lomolino et al. 2013; Heaney 1978).

Given the importance of body size in all aspects of an organism's biology (Schmidt-Nielsen 1984; Peters 1983), size changes are expected to be multifactorial and subject to contingency. But explanations also depend upon what is being explained. In the case of insular body-size change, we must ask, what is the main variable of interest? Is it the direction of size change on an island—smaller or larger, regardless of degree? If magnitude of the change is important, is each increment of difference in a size ratio—say 1.05 to 1.06—biologically equivalent for every set of species, or within some specified groups of species? Is a change from 0.99 to 1.00 truly comparable to a change from 1.00 to 1.01? And if some differences are too small to consider, what is the threshold for what constitutes "meaningful" change?

In this paper we focus on rodents, the most species-rich order of mammals and the order that has presented the greatest obstacle to formulation of a general "rule" for mammalian body size change on islands. A widespread but phylogenetically circumscribed group of mammals, rodents share a distinctive and well-defined set of morphological traits. Yet despite their morphological coherence, rodents vary widely in body size, habitat, and diet. Extant insular populations of rodents exhibit higher

variance in and a wider range of size ratios than any other order, and both the direction and the magnitude of size change can vary among different island populations within a single rodent species (Meiri, Cooper, and Purvis 2008; Durst and Roth 2012; McClain et al. 2013).

Here we address the following questions:

(i) for discerning any "rules" governing body size evolution in insular rodents, what constitutes a meaningful difference in body sizes--that is, size variation that extends beyond that commonly found among populations on the mainland?

(ii) what factors are associated with such extreme—as opposed to more moderate--cases of insular body size change?

(iii) within the set of cases that have shown extreme change, what factors differentiate extremely small- and extremely large-bodied populations?

To address question 1, we use a boot-strapped Kolmogorov-Smirnov (K-S) test to compare the distribution of island:mainland size ratios for 300 insular rodent populations (more specifically, average body size for each island population divided by the average size for its species on the mainland) with the distribution of over 1,000 analogous ratios from the mainland (ratios between the average for a mainland rodent population and its species average). Ratios falling in the tails of the mainland distribution--below the 2.5% or above the 97.5% quantile--are recognized as extreme. For question 2, we examine differences (again using K-S tests) in the distributions of 19

organismal, climatic, and geographic variables between moderate- and extreme-sized insular rodent populations. Finally, for question 3, we use classification trees and random forest techniques to develop a predictive model describing factors associated with extremely large and extremely small body sizes in insular populations. We then test this model on a set of cases independent from the sample used to create it.

2.2 Methods

2.2.1 Island Data Collection

Records of insular rodent populations were taken from the database of McClain et al. (2013), which is an expansion of the database from Meiri et al. (2008). Information for each population included organismal traits (island:mainland body size ratio, mass, dietary data, substrate preference), six climatic variables (temperature and precipitation, each averaged spatially across the island, their standard deviations and their temporal variation), four measures of island productivity (the spatial mean, maximum, minimum, and standard deviation of net primary productivity, NPP) and six geographic variables (measures of island area, elevation, and isolation). Details on data collection can be found in the supplemental material for McClain et al. (2013). Since island area and elevation can both be indicators of island terrain heterogeneity (Mark V Lomolino et al. 2012), we combined measures of the two in a principle components analysis and used PC1 as a measure of island heterogeneity (SI Table S1). The total of 306 populations of

insular rodents (67 species on 182 islands across the world) were included in the analyses.

2.2.2 Mainland Population Collection

We obtained mainland rodent records from the databases at the Natural History Museum of Los Angeles County and the Smithsonian, U.S. National Museum of Natural History. We discarded all records that were missing body measurement or locality information, and then grouped the records together by species and population. Populations were defined as more than one record of the same species obtained from exactly the same location as noted in the museum records. This process resulted in a total of 1,076 rodent populations (mean=5.4 records/population, median=3 records/population) from across the world including species from every family that is represented in the island database. We took the average mass for each population, and compared it to the average mass for the species as defined in the masses of mammals database (Smith et al. 2003), to obtain a population:species body size ratio for each population analogous to the size ratios between island and mainland populations described above.

2.2.3 Defining 'Meaningful' Size Change

To test for a difference between the size ratio distributions of mainland and island populations (each standardized, because sample sizes differed, to total 100% in a probability density function), we used the *Matching* package in R (R Development Core

Team 2013; Sekhon 2011) to perform a bootstrapped Kolmogorov-Smirnov (K-S) test on the two groups.

We then defined 'Normal' size differences for the mainland as any size ratio that fell between the lowest and highest 2.5% of all mainland size ratios. Anything falling outside those boundaries (in the tails) was defined as 'Extreme' size difference. The same numerical size-ratio cutoffs (values of the ratios, as opposed to percentage) were then applied to the island populations to identify insular size ratios that were 'normal' and 'extreme' with respect to variation on the mainland. The extreme island populations were further categorized as either extremely big ('Big') or extremely small ('Small').

2.2.4 Comparing Conditions associated with 'normal' vs 'extreme' Island Populations

To examine the factors influencing extreme insular size change, we plotted as probability density functions the frequency distributions for the 'normal' and 'extreme' cases with respect to each of the geographic, climatic and productivity variables, and then compared these distributions using bootstrapped K-S tests. The pairwise comparisons included 'normal' vs Big, 'normal' vs Small, 'normal' vs Big + Small combined, and Big vs Small. In all instances the threshold for significance was taken as $p < 0.05$. No correction was made for multiple comparisons, as we viewed these comparisons as descriptive and exploratory.

2.2.5 Classification Tree Analyses

In addition to comparing frequency distributions for each variable, we used classification trees to create a predictive model for the directionality of extreme insular size change (Big or Small). CART (classification and regression tree, jointly called "decision tree") methods identify a succession of predictor variables to group samples so as to minimize within-group heterogeneity while maximizing the between-group heterogeneity of a response variable. Both classification and regression trees have been implemented with some success in previous studies, by the current authors and others, on insular mammalian body size change. See chapter one for more extensive discussions of CART methods.

For the island rodent samples deemed 'Extreme' according to the criteria noted above, we constructed a fully-grown classification tree using the *rpart* package in R (Therneau and Atkinson 1997; R Development Core Team 2013) using all of the variables in our insular population dataset to predict the direction of extreme size change. The fully grown tree was pruned to obtain the optimal tree (cross-validated error rate within 1 SE of minimum error tree; as discussed in ref. 5). To test the stability of our classification tree, we also built a random forest model consisting of 5,000 trees built from bootstrapped subsets of the data using the *randomForest* package in R (Liaw and Wiener 2002). A relatively stable classification tree will have a predictive accuracy similar to that of the related random forest model.

Earlier studies that applied CART methods to identify variables associated with insular body size change (Durst and Roth 2012; Lomolino et al. 2012) were largely descriptive and inferential, constructing the decision trees and examining the chosen variables for their biological relevance. However, CART methods, especially classification trees, are widely used in other disciplines not only to describe the data used to build the tree, but to build a model to predict values of the response variable for samples independent of those used to construct the tree (Prasad, Iverson, and Liaw 2006). Recent work by Lomolino et al. (2013) provided size ratios and locality data for almost 100 insular rodent populations for which we had island information (obtained earlier for purposes of studying insular mammals other than rodents; McClain et al. 2013) but had not included in our rodent data set. Of those populations, 39 of them qualified as extreme according to our definition. We ran these 39 samples through our classification tree and random forest to assess the efficacy of our models as predictive tools.

2.3 Results

2.3.1 Defining 'Meaningful' Size Change

According to the K-S test, size ratio distributions for island and mainland rodent populations differed significantly ($p<0.001$; see Figure 4a). On the mainland, the lower 2.5% quantile for size ratios was 0.80 while the upper 97.5% quantile was 1.09. We took these as the boundaries for 'normal' body size, and the 5% of mainland samples falling

outside these boundaries were considered ‘extreme’ sized mainland populations (see Figure 4a). Applying the same boundaries to the island samples, 54% of the samples were classified as ‘extreme’ sized insular populations. Of these extreme sized populations, 78% (deemed 'Big') had increased in size, while 22% (deemed 'Small') decreased in size. Although there was no phylogenetic structuring, above or within species, in the tendency to produce insular populations of extreme body size (see Appendix C), for species that produced extreme populations on multiple islands the tendency was strong for those extreme changes to occur in the same direction (Big or Small).

2.3.2 Comparing Normal and Extreme Islands

The distributions of seven variables were found to differ significantly between insular populations of Big and 'normal'-sized rodents (see Figure 4b-g). Among these seven variables, values were generally lower for heterogeneity of island terrain (a composite measure of area and elevation), mean temperature (averaged spatially), and the spatial standard deviation in annual precipitation for the Big populations, while distance to continent, minimum net primary productivity (NPP), and mean annual precipitation tended to be higher. Three variables were distributed significantly differently between island populations of 'normal'-sized rodents and Small insular populations, which, by contrast, tended to experience high mean annual temperature, a high standard deviation of NPP, and lower mean precipitation.

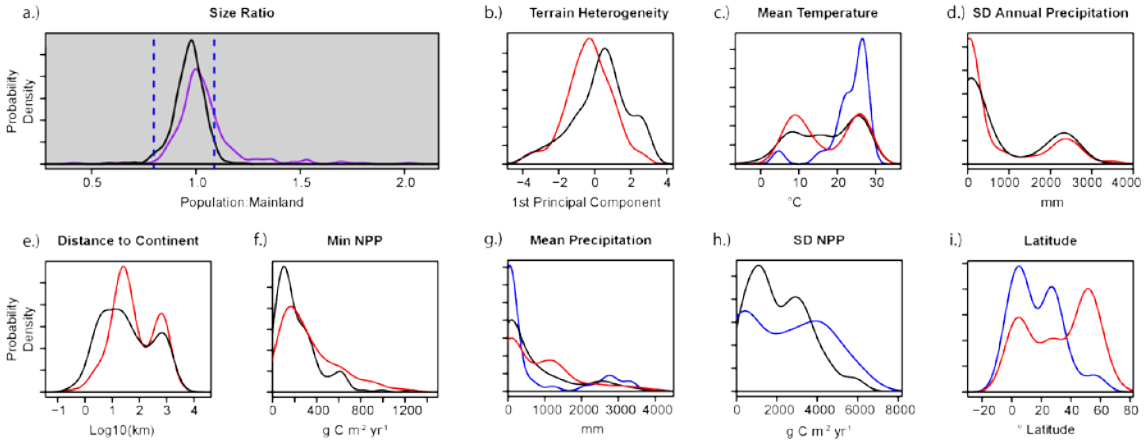


Figure 4. Comparing distributions of numbers of island populations with various characteristics. (a) Size ratios for island (purple) and mainland (black) populations; vertical dashed lines delimit tails of the distributions showing extreme size reduction (left) or increase (right). (b-i) Various attributes whose distributions differ significantly (K-S test, p<0.05) between 'normal' (black), Small (blue) and/or Big (red) island populations. For all graphs, the y axis represents the number of populations at a given value on the x axis.

Big and Small insular populations differed significantly in four variables. Mean precipitation was higher for Big populations, mean temperature and the spatial standard deviation of NPP were both higher for Small populations, and Big populations were found at higher latitudes than Small ones.

2.3.3 Classification Tree Analyses

The classification tree based on our original dataset (Figure 5a) correctly classified 90% of its extreme cases of insular body size changes (98% of the Big and 62% of the Small populations correctly predicted, Cohen's $\kappa = 0.671$, $P < 0.001$) while the random forest correctly predicted 87% of the size changes (97% of the Big and 52% of the Small, Cohen's $\kappa = 0.559$, $P < 0.001$). The first split on the classification tree distinguishes

between different dietary categories. While populations in most diet categories overwhelmingly tend to become Big on islands (102 populations becoming Big, 10 becoming Small), frugivores and granivores vary (30 Big, 27 Small). Size change in frugivores and granivores appears to depend on the mean productivity of the islands

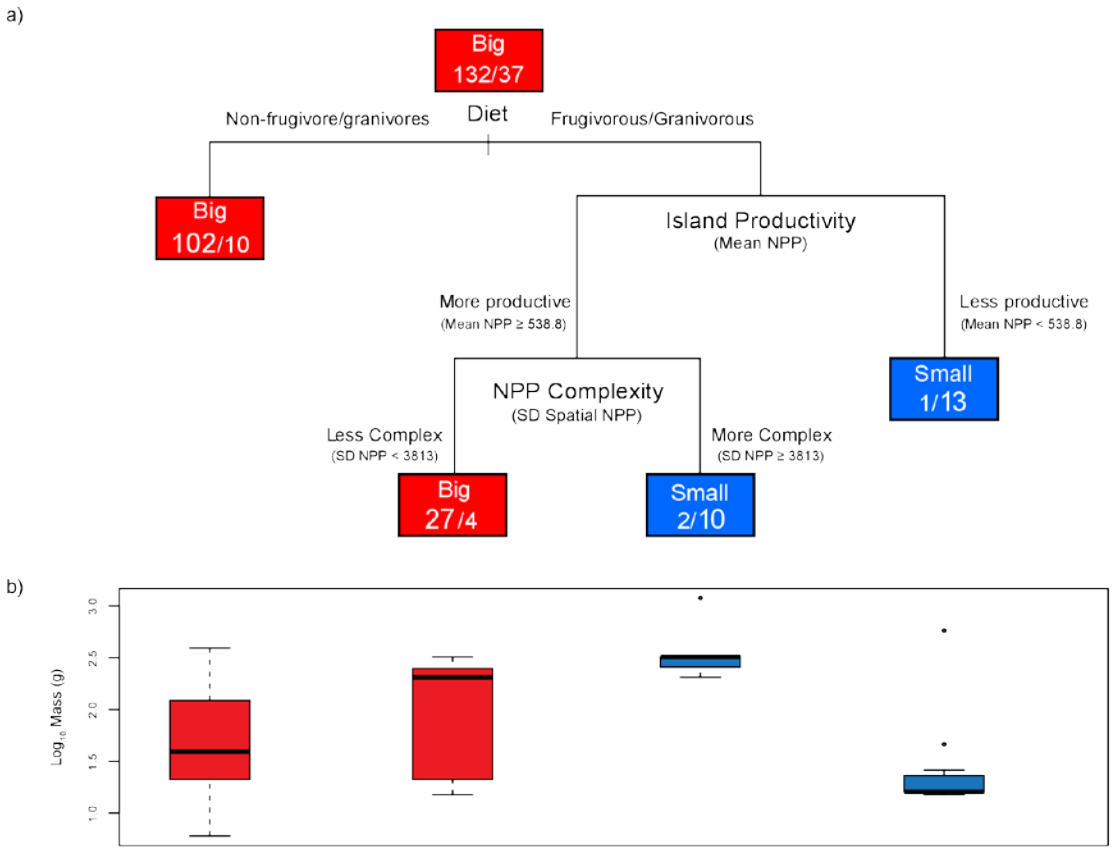


Figure 5. Classification tree, showing predicted direction of size change for insular rodent populations, and body size ranges corresponding to its terminal nodes. (a.) Classification tree showing variables and the ranges of values used in the classification process and the numbers of cases showing size increase / decrease at each node. In the terminal nodes, sets/groups of cases predicted to have increased in size are shown in red; predicted to have reduced size in blue, and numbers of cases predicted correctly are in larger font. (b.) Box-and-whisker plot showing the distribution of species masses (on the y axis) for each of the four terminal nodes in the tree. The box extends from the first to the third quartile with the bold line indicating the median mass for the box.

they inhabit, the tendency being to become Small on low productivity islands (1 Big, 13 Small populations). Frugivore and granivore populations on high productivity islands are subdivided a final time on the basis of the heterogeneity of the productivity landscape on the island. Islands that vary in productivity spatially (i.e., with high spatial productivity standard deviations) tended to produce Small body sizes (2 Big, 10 Small populations) whereas low standard deviations produced Big body sizes (27 Big, 4 Small).

When applied to a new dataset derived from ref. 8, the classification tree was able to correctly classify 33 of the 39 'extreme' populations (85%: 94% of Big populations, 20% of Small populations correctly predicted) while the random forest correctly predicted 90% of the new populations (100% of Big populations, 20% of Small populations correctly predicted).

2.4 Discussion

From this examination of extreme insular body size evolution, it is clear that island rodents exhibit some striking differences from their mainland counterparts. Not only do the distributions of size ratios differ between island and mainland populations, but over half of all insular rodent populations fall outside the boundaries encompassing 95% of the size variation ordinarily found among mainland populations.

We found a greater than expected proportion of insular rodent populations of extremely small body size (12% of all island populations, 4.8 times the frequency

observed on the mainland), but the great majority of extreme sized island rodents were extremely big (42% of all island populations, 16.8 times the frequency on the mainland). These percentages stand in stark contrast to the broader pool encompassing all island rodents in our dataset where it is a virtual coin toss whether a population will increase (to some degree) or decrease in size (58% are relatively big, 42% relatively small). This strong trend for 'extreme' populations to be of larger body size suggests that this is the general pattern for rodents on islands, whereas populations that become extremely small represent notable exceptions. (This tendency is evident even if cutoff points for Big and Small are chosen to be symmetrical about a value of 1; see SI Appendix S2.)

The idiosyncratic nature of insular rodent populations of extremely reduced body size is also reflected in the groups identified by the classification tree. Figure 5b shows the differences in body mass between sets of populations predicted to be extremely big and others predicted to be extremely small. The two terminal nodes in which body size was predicted to be extremely big comprise a large number of populations (112 and 31, respectively, together representing 84.7% of all extreme populations) and have a wide range of masses, encompassing almost the entire range of rodent body sizes considered in this analysis (6-304 grams; populations for which the predictions of size increase were accurate spanned the same range). In contrast, the terminal nodes for which size was predicted to have decreased consist of fewer samples (12 and 14, 15.3% of extreme populations) and have much narrower body size ranges

that are also quite different from each other. In one set, and noting the cases in which the prediction of size decrease was accurate, all but two of the populations of large rodents (mean=357 grams, median=322 grams) had become extremely smaller (comparatively "Small") on islands; in the other set, 13 out of 14 populations of small rodents (mean=24 grams, median=16 grams) had become extremely small. Based on these patterns, it appears that rodents on islands are generally extremely big; that extreme insular size decreases represent a deviation from the more typical pattern, and that each group of size-reduced insular populations requires a distinct set of variables to explain its deviation.

The factors identified as significant by the Kolmogorov-Smirnov tests point to resource-related variables being major drivers of extreme size change. Terrain heterogeneity, temperature, precipitation, and productivity differ in their distributions for islands that produce 'extreme' rather than 'normal' body sizes in their rodent inhabitants: Greatly enlarged rodents tend to be found on small, homogeneous islands with above-average precipitation and high, relatively consistent resource availability; extremely reduced rodents are typically found on hot, dry islands where availability of resources is spatially variable.

On the classification tree, populations of extremely reduced rodents were dispersed among several terminal nodes, suggesting that different insular circumstances lead distinct subgroups to evolve substantially smaller size. Yet despite differences

among these subgroups in the specific sets of relevant variables, resource availability may be an underlying factor that influences all of them. All groups that were successfully predicted to be extremely reduced in size are fruit and grain specialists found on islands with either low or spatially heterogeneous ('complex') productivity, whereas the fruit and grain specialists that were predictably greatly enlarged are found on islands with homogeneous productivity landscapes. Since resource availability declines much faster with decreasing area for fruit/grain specialists than for more generalist species (Lawlor 1982; McNab 1963), small frugivore/granivores on low productivity islands may not be able acquire the resources necessary for maintaining a 'normal' size. Likewise, large frugivore/granivores on islands with a heterogeneous productivity landscape may be similarly limited due to their inability to access patchy resources across an island.

While the classification tree and random forest both excelled in classifying the data used to build the models, their success at predicting independent data is perhaps even more impressive. Of the 34 cases of insular rodent populations in the independent dataset that exhibited extreme size increase, the classification tree successfully predicted all but one while the random forest correctly classified all of them. Although the classification rate for populations of small body size from that dataset was less successful (1 of 5 correctly classified), all four of the misclassified populations reached the same terminal node as the largest group of misclassified populations from the

original data: extremely reduced non-frugivorous/granivorous rodents. Since measurements such as net primary productivity and diet are imprecise proxies for factors like resource availability, we might anticipate that predictions would improve with the inclusion of more detailed diet and resource availability data.

The results of our analyses underscore the importance of taking variation that is commonly found on the mainland into account when identifying factors that produce unusual body sizes on islands. Striking differences exist between mainland and island populations, but many of the factors associated with changes in body size in insular rodents come into focus only when the statistical noise produced by variation ordinarily found also on the mainland is removed. Specifically, it becomes clear that some version of the island rule holds true for rodents, with the large majority of insular populations that undergo extreme changes increasing in size. While exceptions to this rule still exist, we have developed a predictive model based largely on diet and resource availability that succeeds in classifying over 90% of extreme populations, remains robust when classifying novel samples, and suggests avenues for future fine-tuning. Although previous authors have despaired of finding generalities that could predict the response of populations to novel insular environments, using models designed to handle the very factors and contingencies that frustrated previous work, we find that significant patterns do emerge. These patterns and the resulting predictability of size change found in this

study represent a major step in the description and understanding of an "ecotypic rule" half a century in the making.

2.5 Reflections on Chapters One and Two

It has been 50 years since Foster published his paper documenting the trends in mammalian insular body size that have come to be known as the island rule, the tendency for insular mammals to be different in size from their mainland counterparts, yet it continues to be a topic that generates scientific debate. A review of the island rule literature since then reveals two distinct types of studies in terms of methods and results. One class focuses on small taxonomic groups (generally a single species or genus) and finds generally strong support for a few variables leading to insular body size change (Adler and Levins 1994; Palmer 2002; Bromham and Cardillo 2007; Benton et al. 2010). The second class analyzes a much wider set of species and factors and finds weak or no support (Lomolino 1985; Lomolino 2005; Meiri, Dayan, and Simberloff 2006; Meiri, Cooper, and Purvis 2008; Lomolino et al. 2012; McClain et al. 2013). Despite these large-scale findings, the striking examples of insular body size change along with the sheer number of documented cases of insular body size change, including the repeated size change of a single species on multiple islands, are difficult to ignore.

For my work on insular body size change, I set out to find a middle ground between the singularly-focused and the all-encompassing studies. Focusing on rodents allowed me to work with a large taxonomic group, but one where some ecological

generalities (life history traits, diet, etc…) still existed. I also tried to explicitly acknowledge what I found to be one of the bigger weaknesses of previous studies of insular body size: the issue of time. The length of time since a species colonized an island can have a large impact on the degree of size change observed, but time since colonization is very difficult to estimate without molecular data, well-dated fossils, or relavent historical records. The scarcity of this type of data makes large-scale studies nearly impossible. The first two chapters deal with the problem of time somewhat differently. In the first chapter, the focus is on the direction of size change, independent of magnitude. The second chapter, only analyzed populations that had undergone extreme size change. While both of these methods have their limitations (analyzing the direction of change assumes that all size change, regardless of degree, is meaningful; only looking at extreme populations may eliminate potentially informative populations that are not yet 'extreme'), they allowed me to focus on factors influencing size change without conflating them with the mechanisms required to enact these changes.

Despite the very different approaches taken in the two chapters, both studies yielded mutually consistent results. While species-level traits are important in determining direction and degree of size change, traits for higher taxonomic levels seem to have relatively little impact. Island traits are highly predictive of insular body size change, reinforcing the idea that island environments have unique features that lead to extreme size change not seen on the mainland. Both studies found that successful

predictions of insular body size change do not rely on ecological interactions with other species, and both studies suggest that resource availability (whether in terms of dietary range, island productivity, or the accessibility of resources within an environment) is one of the major drivers of insular body size change. Perhaps most importantly, these studies provide a link between broad-scale analyses and single-species studies by examining a diverse sample of insular rodents and producing hypotheses that can be explicitly tested in a more limited framework with smaller taxonomic groups, where fossils, historical records, and molecular data can more easily be integrated into the analyses. While the debate rages on about what the island rule actually entails, the fact that predictable patterns can be gleaned from a group as large and varied as rodents provides hope that generalities may exist and that we may one day understand the mechanisms producing the weird and wonderful size changes that have captured the imagination of scientists for 50 years.

3. A mitochondrial phylogeography of *Peromyscus maniculatus* across the northern California Channel Islands

3.1 Introduction

The California Channel Islands, a collection of eight islands off the coast of southern California, are one of the only coast Mediterranean-type ecosystems in North America and as such, they harbor a unique suite of endemic plants and animals. At least six mammals are endemic to the islands, with four of them found on multiple islands (the island fox, spotted skunk, deer mouse and harvest mouse; Floyd et al. 2011), but the deer mouse, *Peromyscus maniculatus*, is the only mammal found on all eight of the islands, with each island housing a morphologically and genetically distinct subspecies (Pergams, Lacy, and Ashley 2000). A great deal of work has been done on the morphology, development and behavior of Channel Islands deer mice (Roth and Dawson 1996; Roth and Klein 1986; Orrock 2010), but the genetics of the species remains unclear.

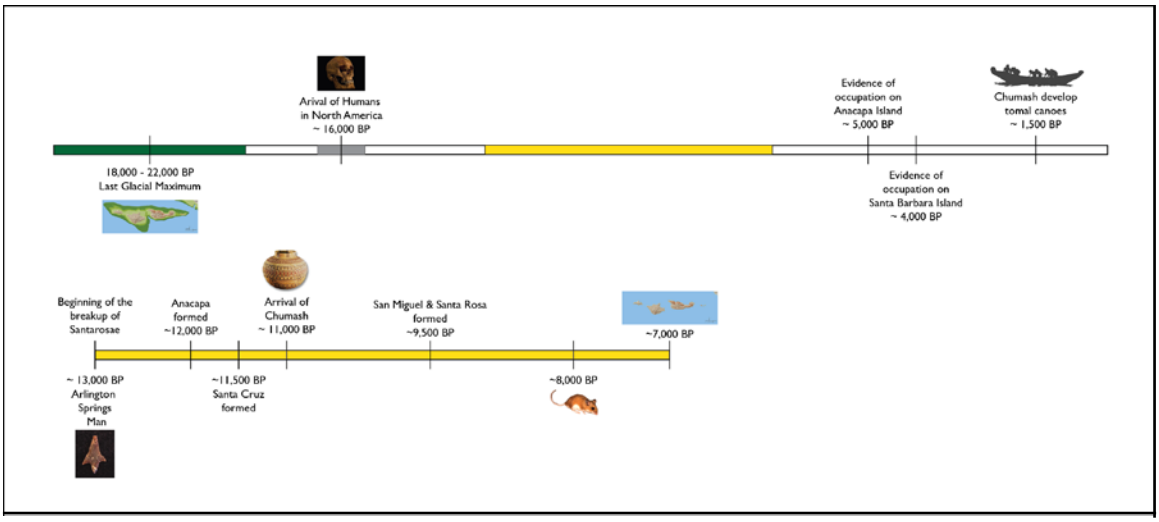


Figure 6. Timelines of important events in Channel Islands history.

To understand the distribution of *P.maniculatus* across the Channel Islands, it is important to understand the geography and geological history of the islands. The Channel Islands are broken into four northern islands (Anacapa, Santa Cruz, Santa Rosa and San Miguel) and four southern islands (Santa Barbara, San Nicolas, Santa Catalina and San Clemente; Figure 7). The four southern islands have been isolated for millions of years, having never been connected to other islands or to the mainland. While the four northern islands have never been connected to the mainland, they were connected to each other during the last glacial maximum (~18,000 years ago), forming the larger island of Santarosae. Bathymetric measurements suggest that the breakup of Santarosae began with Anacapa splitting off 12,000 years ago. Santa Cruz followed 11,500 years ago, and San Miguel and Santa Rosa split 9,500 years ago (Floyd et al. 2011).

Because the earliest fossils of *P.maniculatus* on the Channel Islands date to approximately 8,000 years ago (Guthrie 1993), some have argued that deer mice arrived on the islands after the breakup of Santarosae. This is in line with findings for other mammals, such as the island fox where radiocarbon dating of the earliest known island fox remains postdates the colonization of the islands by ancient humans (~12,000 years, Erlandson et al. 2011). However, allozyme, RFLP and morphological data for the deer mice all suggest that, at the very least, the three large northern island populations (Santa Cruz, Santa Rosa and San Miguel) are more related to each other than they are to mainland mice (Gill 1976; Ashley and Wills 1987), leading others to argue for an arrival prior to the breakup of Santarosae. In their study examining 9 restriction fragments, Ashley and Wills even went as far as to present possible routes of colonization for each island population (Figure 7). While San Clemente and Santa Catalina had unique haplotypes most similar to the mainland to the east, Ashley and Wills found shared patterns between three of the northern islands (Santa Cruz, Santa Rosa and San Miguel) and the mainland to the north, suggesting a colonization of those islands from the northern mainland. They found similar patterns on the islands of Santa Barbara and San

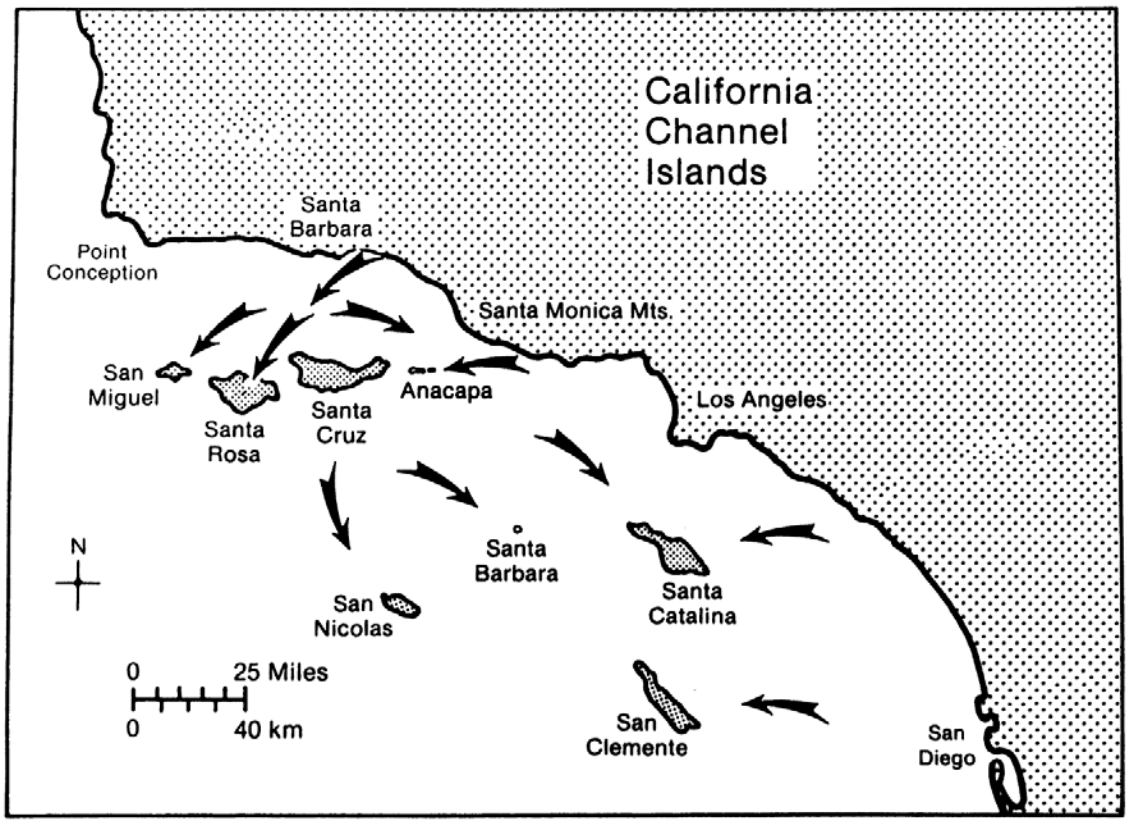


Figure 7. Map of the California Channel Islands. Arrows indicate the colonization routes proposed by Ashley and Wills (adapted from Ashley and Wills, 1987).

Nicolas, which they suggested could be due to a later colonization of those islands from the northern islands. Additionally, they found evidence for a secondary introduction of deer mice to Anacapa Island from the eastern mainland.

The work by Ashley and Wills over thirty years ago was the last multi-island molecular study of Channel Islands deer mice. Due to the limited inferences that can be made from RFLP data (RFLP analyses are unable to identify specific base pair changes in

the genome, only tracking changes in the overall size of different restriction enzyme fragments), a great deal of uncertainty remains regarding the history of *Peromyscus* across these islands. In this study, I use mitochondrial sequence data to reassess the patterns described by Ashley and Wills, and to create an updated phylogeography of *Peromyscus maniculatus* across the northern Channel Islands.

3.2 Methods

3.2.1 Sampling

A total of 72 *P. maniculatus* individuals were collected across six sites on the northern Channel Islands: one site on Anacapa, Santa Rosa, San Miguel and Santa Barbara Islands, and one site each on both the National Park and Nature Conservancy sides of Santa Cruz Island (referred to as Santa Cruz East and West, respectively). An additional 21 mice were collected from two mainland sites: one site in the Santa Ynez Mountains (north of the city of Santa Barbara) at Sedgwick Reserve and one site at Grant Park, located in the city of Ventura. These mainland sites were chosen to examine the colonization routes hypothesized by Ashley and Wills (1987), with Sedgwick representing a northern colonization route, and Grant Park a colonization route from the east.

3.2.2 Molecular Genetics

Whole genomic DNA was extracted from all individuals using a DNEasy Blood and Tissue Kit (Qiagen). A ~400 bp region of the mitochondrial control region was

amplified and sequenced using the PeromtD-F4 (5′-TCTGGTTCTTACTTCAGGGCC-3′) and PeromtD-R (5′-GCATTTTCAGTGCTTTGCTTTATTG-3′) primers. The polymerase chain reaction cycle for the amplification consisted of 94 °C for 2 min; 40 cycles of 94 °C for 15 s, 50 °C for 15 s and 68 °C for 1 min; 68 °C for 10 min. Reactions were cleaned using ExoSAP and sequenced using the BigDye Terminator Cycle Sequencing system run on a 3730XL Genetic Analyzer (Applied Biosystems Inc.). Forward and reverse sequences were assembled using Sequencher v4.8 (Gene Codes) and aligned using ClustalX2 (Larkin et al. 2007). All sequences will be deposited in GenBank concurrent with the publication of this work.

3.2.3 Population Genetic, Phylogeographic and Phylogenetic Analyses

Nucleotide diversity (π), pairwise F_{ST} values, and Tajima's D and Fu's FS tests were all calculated for each population using Arlequin v3.5 (Excoffier and Lischer 2005). I then split the populations into two regions, island and mainland samples, and performed the same analyses, for these groups. Significance values were adjusted for multiple comparisons using the Bonferroni correction (Rice 1989). I also calculated the site frequency spectrum for each group using the PEGAS package in R (Paradis 2010) to look at differences in haplotype diversity between island and mainland populations. To examine the importance of within-island, within-region and between-region genetic variation, I performed an AMOVA (also in Arlequin). Significance was assessed using a permutation procedure (10,000 replicates).

While traditional population genetics parameters are able to provide information about differences in allele frequencies between populations, they do not offer much insight into the relationships between haplotypes and populations. To visualize this relationship in a phylogeographic framework, I constructed a statistical parsimony network as implemented in TCS v1.21 (Clement, Posada, and Crandall 2000). In the few instances where the network resulted in ambiguous loops, I resolved the ambiguities following criteria predicted by coalescent theory as described by Uthicke and Benzie (2003).

For the phylogenetic analyses, sequences were analyzed by PAUP* v4.0 (Swofford 2002) using the MrModelTest block. These scores were submitted to MrModelTest v2.3 (Nylander 2004) to select the best model of evolution for the data (HKY+G). While previous phylogeographic studies of *Peromyscus maniculatus* in other regions have been able to take advantage of traditional phylogenetic frameworks (Yang and Kenagy 2009; Taylor and Hoffman 2010), the recent colonization history of Channel Islands *Peromyscus* meant that traditional maximum likelihood and Bayesian phylogenetic methods yielded relatively uninformative and unsupported phylogenies for my dataset. As such, analyses of Channel Islands *Peromyscus* were conducted in BEAST v1.8 (Drummond et al. 2012) using a Bayesian coalescent framework. Using the model suggested by MrModelTest under both a constant size and exponential growth population model, I ran the analyses for 10,000,000 generations, logging every 1,000

steps. The logs were imported into Tracer v1.6 (Drummond and Rambaut 2007) for review. Because the marginal posterior density of the population growth parameter did not overlap zero, I imported the trees generated from the exponential growth population model into TreeAnotator 1.8 (Drummond and Rambaut 2007) to compute the consensus tree.

3.3 Results

I sequenced and aligned 394 bp from the mitochondrial control region for each of the 93 individuals sampled. Overall nucleotide diversity was understandably low (0.6%, 2.35 pairwise nucleotide differences) given the relatively small spatial and temporal scale of this study, but nucleotide diversity was significantly lower on islands (0.45%, 1.76 pairwise differences) than on the mainland (0.75%, 2.94 pairwise differences, p=0.020 per

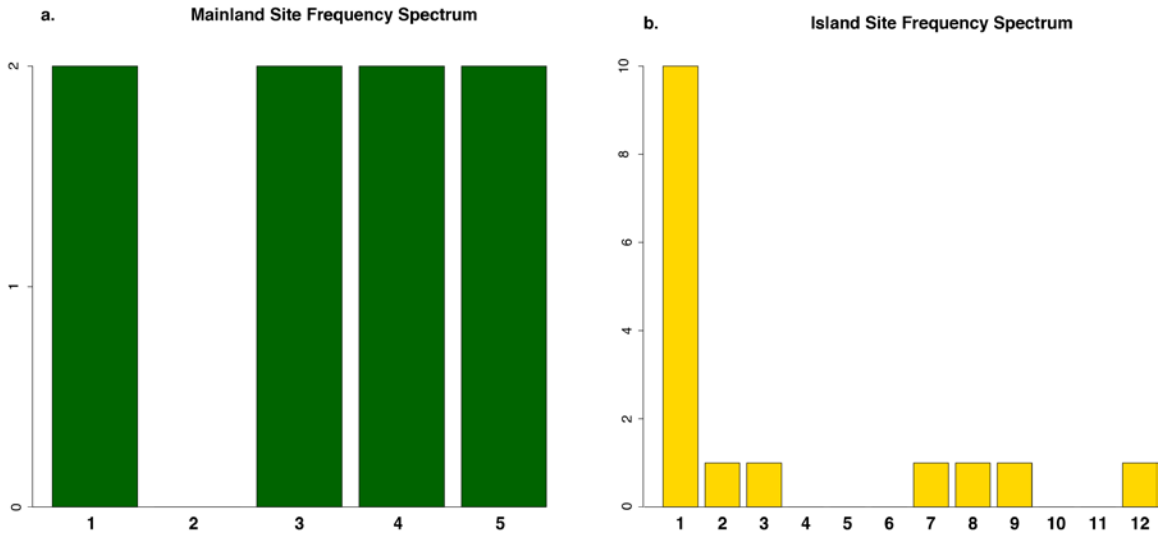


Figure 8. Site frequency spectrums for (a.) mainland and (b.) island populations. In this histogram, the *i*th entry (x axis) is the number of polymorphic sites (y axis) for which the mutant allele is present in *i* individuals.

Welch's t-test). The site-frequency spectrum analyses revealed an excess of rare alleles on islands when compared to the pattern on the mainland (Figure 8). Tajima's D values were significant and negative for San Miguel (D=-1.60, p=0.046) and Santa Barbara (D=-1.61, p=0.033) and Santa Barbara was the only population with a significant value for Fu's FS (FS= -3.33, p=0.004). At the group level, both Tajima's D and Fu's FS were significantly negative for the island group (D=-1.76, p=0.016, FS=-14.34, p<<0.001). Pairwise F_{ST} comparisons revealed significant differentiation between all populations with the exception of the northern islands of Santa Cruz, Santa Rosa and San Miguel (Table 1). Inter-island F_{ST} values between Anacapa and other islands were more similar to F_{ST} values found between island and mainland sites (F_{ST} = 0.45-0.60). Inter-island F_{ST} values for Santa Barbara Island were also consistently higher than other inter-island values (F_{ST} = 0.368-0.509). The pairwise F_{ST} for island and mainland groups was similar to that for individual island-mainland comparisons (F_{ST}=0.61, p<<0.001). For the AMOVA, mainland vs. island populations accounted for the largest percentage of variation (57.30%, p<<0.001) and while within-group and within-population differences both accounted for a significant portion of the overall variation, population differentiation within island and mainland groups accounted for the smallest amount of variation (Table 2).

The haplotype network (Figure 9) revealed a single dominant haplotype shared by at least one individual from each of the six island populations along with one

individual from Sedgwick Reserve. This haplotype and a haplotype shared by individuals from Santa Cruz and Santa Rosa were the only haplotypes shared between island populations. Each island had its own suite of unique haplotypes stemming off from the dominant island haplotype, and all mainland haplotypes (other than the shared dominant haplotype) were separated from the dominant island haplotype by at least two base pairs. All unique haplotypes from Santa Barbara Island were connected to each other, stemming from the dominant island haplotype, but there was little other geographic differentiation within either the mainland or island haplotype groups.

The consensus tree from BEAST was in line with the findings from both the population genetics and haplotype network analyses. The only split with 100% posterior support was at the base of the tree between mainland and island haplotypes (Figure 10; the haplotype tree is presented for simplicity's sake). The two exceptions to this island-mainland split are both haplotypes recovered from single individuals at Sedgwick Reserve (the dominant island haplotype and a unique, closely related haplotype). Within the mainland clade, there was little support for geographic structure. The only clade that was somewhat supported (posterior probability = 0.69) consisted of two Grant Park haplotypes and a Sedgwick Reserve haplotype. Within the island clade, the only strongly supported (posterior probability > 0.90) group was a clade consisting of all the unique Santa Barbara island haplotypes. A clade of 3 Santa Rosa haplotypes, and a clade with a pair of Anacapa haplotypes were also somewhat

supported (posterior probability = 0.86, 0.78 respectively) along with a clade of consisting of haplotypes from Santa Cruz East, Santa Cruz West and the only other shared haplotype (between Santa Cruz West and Santa Rosa; posterior probability = 0.65), but no other nodes had posterior probabilities > 0.50.

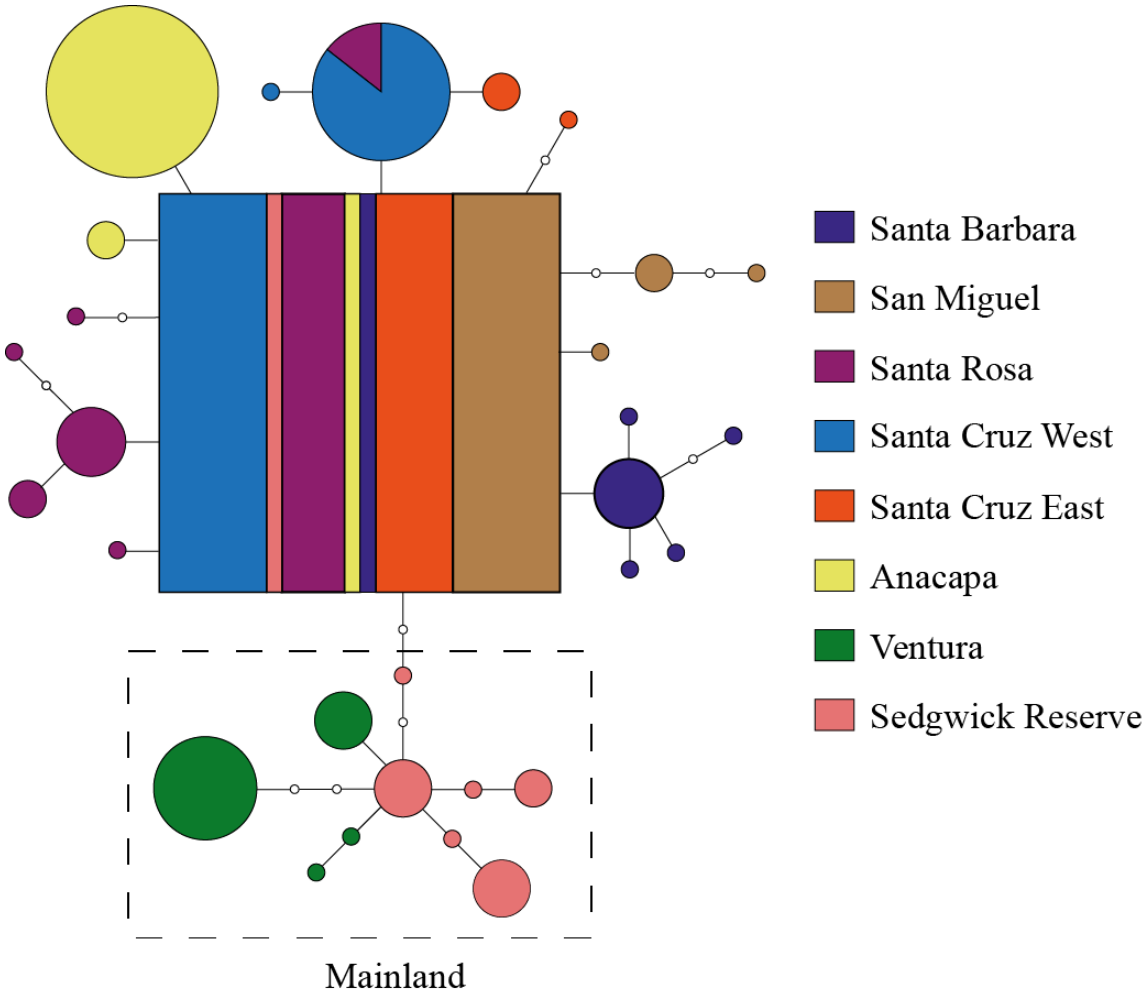


Figure 9. Haplotype network constructed from mitochondrial control region sequence data for northern Channel Island Peromyscus maniculatus and its mainland relatives. Each shape is a single haplotype, and the size of the shape reflects the number of individuals sharing that haplotype. The size of each color within a shape represents the number of individuals from each location exhibiting that haplotype. Small white circles represent unsampled haplotypes.

Table 1. Pairwise FST estimates from the mitochondrial control region for populations in this study.

	Grant Park	Sedgwick Reserve	Anacapa	Santa Cruz East	Santa Cruz West	Santa Rosa	San Miguel	Santa Barbara
Grant Park	0							
Sedgwick Reserve	0.233	0						
Anacapa	0.777	0.675	0					
Santa Cruz East	0.661	0.548	0.509	0				
Santa Cruz West	0.740	0.646	0.640	0.052	0			
Santa Rosa	0.684	0.585	0.487	0.128	0.234	0		
San Miguel	0.684	0.584	0.486	0.109	0.231	0.166	0	
Santa Barbara	0.709	0.633	0.658	0.368	0.509	0.390	0.374	0

Table 2. Differentiation of the populations in this study as determined by an analysis of molecular variance (AMOVA).

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
Mainland vs. Island	1	53.844	1.48352 Va	57.30
Within Mainland or Island	6	27.904	0.35249 Vb	13.61
Within Populations	81	61.005	0.75315 Vc	29.09
Total	88	142.753	2.58917	

Fixation Indices		P values
FSC :	0.31881	<< 0.001
FST :	0.70911	<< 0.001
FCT :	0.57297	0.024

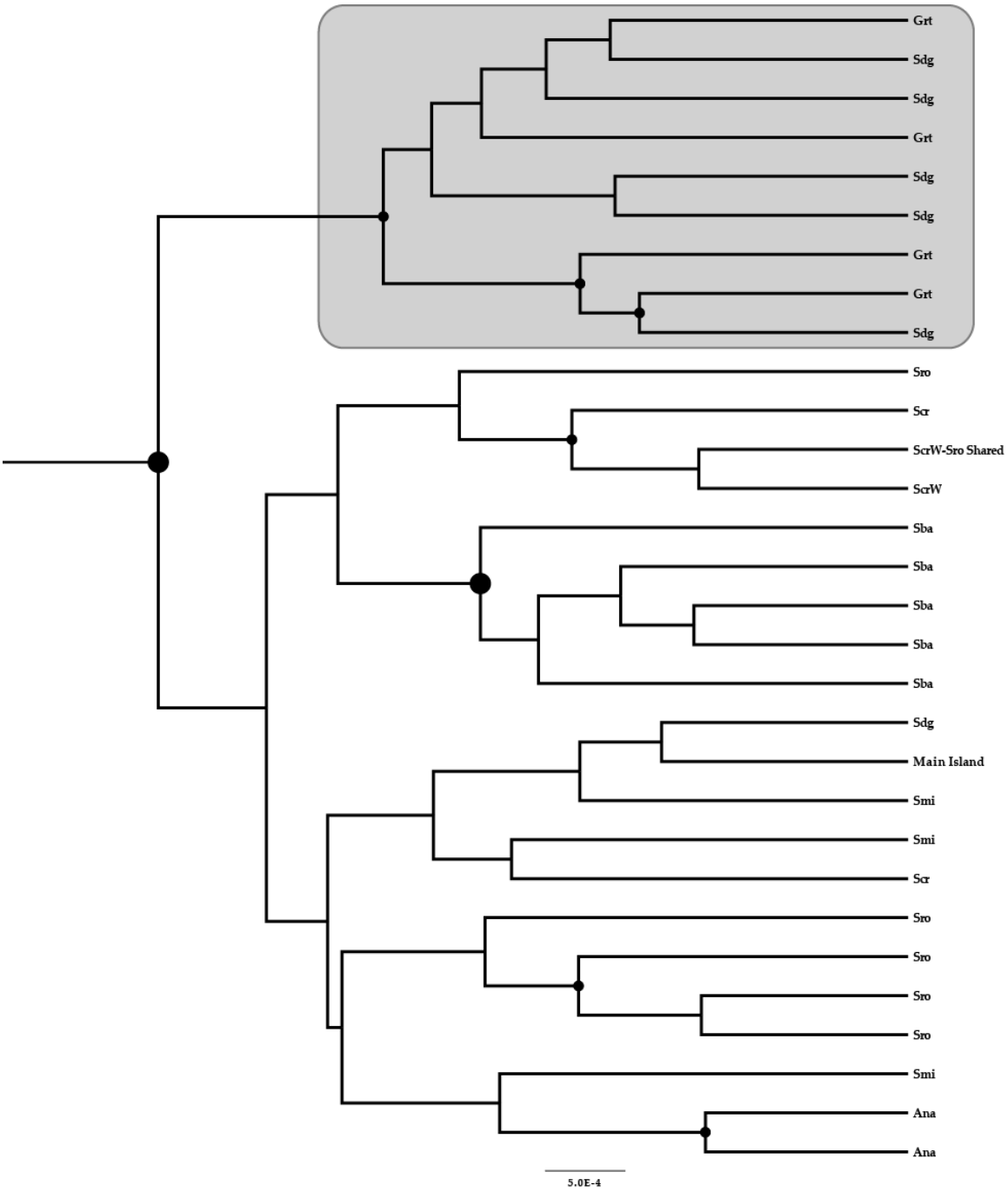


Figure 10. Bayesian phylogenetic tree for the haplotypes of mitochondrial control region sequences present in this study. The grey box outlines the mainland haplotypes. Large circles indicate posterior support greater than 90%. Smaller circles indicate support above 50%. Grt=Grant, Sdg=Sedgwick, Ana=Anacapa, Scr=Santa Cruz, ScrW=Santa Cruz West, Sro= Santa Rosa, Smi=San Miguel, Sba=Santa Barbara

3.4 Discussion

Results from population genetic, phylogenetic and phylogeographic analyses revealed that the most significant differentiation is found between mainland and island populations. The significantly negative values for both Tajima's D and Fu's FS suggest a recent population expansion for the island group, which is consistent with a relatively recent colonization event. Additional support for this conclusion includes the findings that the island-mainland split is the most supported split in the phylogeny, island-mainland differences account for more than half of the variation observed in the AMOVA, and island and mainland populations generally have distinct mitochondrial haplotypes. The finding that Sedgwick Reserve harbors individuals with the dominant island haplotype, along with the two Sedgwick haplotypes nested within the island clade of the phylogeny, supports the hypothesis of a northern colonization of the northern Channel Islands as proposed by Ashley and Wills (1987). Additionally, the fact that all of the northern islands (with the exception of Anacapa; more on that below) are more closely related to each other than they are to either mainland site suggests that deer mice colonized the northen islands prior to the breakup of Santarosae.

The lack of geographic differentiation between most of the northern Channel Islands is not surprising given the relatively short timeframe of this study (< 10,000 years) and the potential for human-mediated migration between the islands. Despite these recent divergences and migration possibilities, I still found significant

differentiation on Santa Barbara and Anacapa islands. The differentiation of the Santa Barbara population in relation to all of the northern islands could represent an earlier divergence of Santa Barbara mice from the Santarosae population before the breakup of the northern islands. Alternatively, this could also suggest that gene flow has continued between the northern islands to a greater extent than it has between the northern islands and Santa Barbara Island. Regardless, the shared haplotypes between Santa Barbara and the northern islands and the similar F_{ST} values between the mainland and all of the islands (including Santa Barbara) suggest, as Ashley and Wills hypothesized, that Santa Barbara was colonized from the northern islands and not from the nearby mainland.

The differences observed on Anacapa Island may have less to do with the breakup of Santarosae than with the recent extirpation of black rats on the island. In 2001-2002, a rodenticide was deposited across of three islets of Anacapa to eliminate an introduced population of black rats that were severely impacting both native mouse and seabird populations. The application of the poison was staggered over two years, with at least one islet remaining rodenticide-free at all times. A survey of Anacapa deer mice before and after the extirpation (Ozer, Gellerman, and Ashley 2011) found that only one of 3 COII haplotypes observed on Anacapa before 2001 survived the extirpation. While the Anacapa deer mice I sampled for this study had a similar number of haplotypes when compared to other islands, the Anacapa population as a whole had the lowest nucleotide diversity (0.12%, 0.47 pairwise differences) of any population. This, along

with higher pairwise F_{ST} values than are seen on other islands, supports the findings of Ozer et al that the Anacapa deer mouse population has undergone a recent loss of diversity, most likely due to the management strategy employed during the extirpation of rats on the island.

This study largely supports the findings of previous researchers, but the work is presented in a more modern framework, integrating population genetic, phylogenetic and phylogeographic methods into a more comprehensive analysis. All of the northern Channel Islands populations appear to share a single origin, potentially from the mainland to the north. Santa Barbara mice are genetically distinct from northern islands mice, but they are more distantly related to mainland mice, suggesting a northern islands origin possibly before the breakup of Santarosae. Deer mice on Anacapa have much lower genetic diversity than mice on other islands, most likely due to the recent extirpation of black rats on the island. While there does not appear to be much support for a secondary introduction of deer mice from the eastern mainland to Anacapa, as Ashley and Wills suggested, the recent loss of genetic diversity may limit our ability to satisfactorily explore this possibility. Over the last 10,000 years, the Channel Islands have been subject to the continued influence of humans. In that time, repeated introductions and manipulations have dramatically altered the landscape of the islands, but *Peromyscus maniculatus* has remained a constant presence on every island. Their history provides an excellent case study for the interaction between natural processes

and anthropogenic influences, and our understanding of these populations will be important for evaluating and managing not only species on the Channel Islands, but rare and threatened species across the world.

4. A comparison between mitochondrial sequence data and microsatellite loci for *Peromyscus maniculatus* on the California Channel Islands

4.1 Introduction

When molecular data is used to infer the evolutionary history of an organism, different genes have the potential to tell remarkably different stories (Nichols 2001). For example, mitochondrial DNA, widely used to study phylogeographic patterns (Avise 2000), is non-recombinant and maternally inherited, resulting in an effective population size for mitochondrial loci four times smaller than the effective population size for nuclear loci. Because of this, mtDNA is much more prone to the effects of genetic drift, reducing problems associated with linage sorting and ancestral polymorphisms (advantageous in a phylogenetic study), but making it harder to observe rare gene flow events between populations (problematic in a phylogeographic study; Ballard and Whitlock 2004). This is especially true for species with male-biased dispersal such as *Peromyscus maniculatus* (King 1968). Several recent regional phylogeographic studies of deer mice using mitochondrial sequences have revealed distinct mtDNA breaks in areas lacking geographic barriers preventing gene flow (Dragoo et al. 2006; Taylor and Hoffman 2010). Subsequent analyses of nuclear microsatellite loci for these populations have found little geographic structure(Yang and Kenagy 2009; Taylor and Hoffman 2011), suggesting that while analyses of mitochondrial sequences may be able to identify

ancient barriers to gene flow, nuclear microsatellite loci may be necessary to capture all of the contemporary gene flow between populations.

In the previous chapter, I found significant differentiation between populations of *P.maniculatus* on the Channel Islands and the nearby mainland. Each island had its own unique mtDNA haplotypes, and all island populations were more similar to other island populations than they were to the mainland populations. The Channel Islands have been physically isolated from each other for 5,000 years (Floyd et al. 2011), theoretically preventing any significant gene flow between islands. However, there has been human activity on the Channel Islands for at least 10,000 years (Erlandson et al. 2011), and based on the findings for other mammals on the Channel Islands (Wayne et al. 1991; Rick et al. 2009; Floyd et al. 2011), the potential for human-mediated gene flow among deer mice across the islands cannot be ignored. In situations like this, it is essential to integrate multiple molecular datasets to better understand the interplay between ancient vicariant events and more modern migration events. Therefore, in this chapter, I present the preliminary findings of an analysis of nuclear microsatellite loci for Channel Islands deer mice. I compare these findings with those from the mitochondrial sequence analyses, and I test for discordance between the two studies.

4.2 Methods

4.2.1 Sampling and Molecular Genetics

For this study, I used the whole genomic DNA extracted from the 93 *P.maniculatus* sampled for the previous chapter. I genotyped 5 microsatellite loci for these individuals using primers from Mullen et al (BW3_15, BW4_8, BW4_12, BW4_7, BW4_200; 2006). Reactions followed the protocol detailed in Mullen et al. (2006). PCR products were analyzed by Eton Biosciences on a 3730XL Genetic Analyzer (Applied Biosystems Inc.) using the LIZ500 size standard. Peaks were identified by hand using the Peak Scanner v1.0 software (Applied Biosystems Inc.), and most homozygous individuals were re-genotyped at least once for quality control purposes.

4.2.2 Microsatellite Analysis

Basic diversity statistics (number of alleles, A; allelic richnes, R_S; number of private alleles; expected heterozygosity, H_E; observed heterozygosity, H_O) were calculated for each population using the PopGenKit package and tests for deviations from Hardy Weinberg equilibrium were conducted using the adegenet package, both implemented in R (Paquette 2012; Jombart 2008). Pairwise F_{ST} values and their associated significance values (determined using 10,000 permutation iterations) were calculated using FSTAT version (Goudet 1995). All significance values accounted for multiple comparisons using a Bonferroni correction (Rice 1989). The matrices of

pairwise F_{ST} values for the mitochondrial and microsatellite datasets were compared using a Mantel test.

To investigate the degree of differentiation between populations, I used the Bayesian clustering program STRUCTURE (Pritchard, Stephens, and Donnelly 2000), assigning individuals to between K=2 to K=8 clusters, the maximum number of clusters corresponding to the number of populations in the study. For each value of K, I ran five replicates with a burn-in period of 10,000 iterations followed by 100,000 MCMC generations. I aggregated these results in STRUCTURE HARVESTER (Earl and vonHoldt 2012), an online web application that consolidates multiple STRUCTURE runs and then calculates the most like value for K using both the log likelihood (ln Pr(*X*/K)) method of Pritchard, Stephens, and Donnelly (2000) as well as the ΔK method recommended by Evanno, Regnaut, and Goudet (2005). The files for the most likely K value were run through CLUMPP (Jakobsson and Rosenberg 2007), which returned average assignments for individuals across all runs, and the CLUMPP files were imported into DISTRUCT (Rosenberg 2004) to create barplots detailing the probabilities of membership in each cluster for each individual.

4.3 Results

4.3.1 Genetic Diversity Analyses

The average number of alleles per locus ranged from 5.0-7.6, while allelic richness (corrected for sample size) ranged from 6.59-7.03 on the mainland and 4.67-6.4

on islands. Each population had at least one private allele, but the mainland averaged 1.6 private alleles, while islands averaged just 0.7. Expected heterozygosity averaged across loci ranged from 0.63-0.80 and observed heterozygosity ranged from 0.56-0.77. All values along with population/locus pairs deviating from Hardy Weinberg equilibrium can be found in Table 3.

4.3.2 Genetic Structure Analysis

Pairwise F_{ST} comparisons revealed significant differentiation between all populations with the exceptions of the two Santa Cruz sites and the Santa Cruz East – Santa Rosa pair (Table 4). A Mantel test comparing the F_{ST} values for the microsatellite data with those from the mitochondrial data found a significant correlation between the two matrices ($p = 0.003$). STRUCTURE HARVESTER determined K=5 to be the most likely number of clusters. Using the STRUCTURE analyses for K=5, both mainland populations form a single cluster, Anacapa, San Miguel and Santa Barbara form their own unique clusters, and Santa Rosa and the two Santa Cruz sites share membership in the Santa Barbara and San Miguel clusters along with a fifth unique cluster (Figure 11).

Table 3. The number of alleles (A), allelic richness (RS), number of private alles (priv), expected heterozygosity (He) and observed heterozygosity (HO) for each population in this study. Values in bold indicate significant deviations from HWE. Starred values indicate values that were significant after a Bonferroni correction.

		BW3_15	BW4_8	BW4_112	BW4_7	BW4_200	Mean
Grant Park	A	7	7	6	7	7	6.8
	R_S	7	6.73	5.43	7	6.8	6.592
	priv	1	3	0	3	1	1.6
	H_E	0.797	0.79	0.745	0.825	0.81	0.793
	H_O	**0.333***	**0.8**	**0.8***	0.8	**0.6**	0.733
Sedgwick Reserve	A	10	7	7	5	8	7.4
	R_S	9.16	6.8	6.18	5	8	7.028
	priv	1	0	1	1	5	1.6
	H_E	0.893	0.79	0.765	0.735	0.808	0.798
	H_O	0.583	0.8	0.8	0.889	**0.444***	0.768
Anacapa	A	4	6	9	4	9	6.4
	R_S	3.92	5.76	7.06	3.35	7.97	5.612
	priv	0	0	0	2	2	0.8
	H_E	0.66	0.766	0.832	0.57	0.799	0.725
	H_O	0.692	0.846	**0.615**	**0.692***	0.769	0.731
Santa Cruz East	A	9	4	7	2	10	6.4
	R_S	9		7	2	10	6.4
	priv	0	0	1	0	2	0.6
	H_E	0.845	0.512	0.836	0.401	0.87	0.693
	H_O	**0.333***	0.667	0.875	0.333	**0.778***	0.625

		BW3_15	BW4_8	BW4_112	BW4_7	BW4_200	Mean
Santa Cruz West	A	7	3	6	2	12	6
	R_S	6.25	2.96	5.37	2	10.15	5.346
	priv	1	0	0	0	2	0.6
	H_E	0.834	0.538	0.768	0.294	0.898	0.666
	H_O	**0.357**	0.714	0.714	0.357	**0.929**	0.614
Santa Rosa	A	11	6	7	2	12	7.6
	R_S	9.1	5.24	5.51	2	9.76	6.322
	priv	2	0	1	0	3	1.2
	H_E	0.879	0.768	0.77	0.477	0.886	0.756
	H_O	**0.385***	**0.5***	**0.286***	0.357	0.857	0.607
San Miguel	A	9	2	6	2	7	5.2
	R_S	7.45	2	5.53	2	6.39	4.674
	priv	3	0	0	0	0	0.6
	H_E	0.809	0.444	0.785	0.413	0.701	0.63
	H_O	**0.667***	0.167	0.833	0.417	0.833	0.563
Santa Barbara	A	7	3	7	3	5	5
	R_S	7	3	6.91	3	5	4.982
	priv	0	0	1	1	0	0.4
	H_E	0.741	0.426	0.845	0.586	0.777	0.675
	H_O	**0.556***	0.333	1	0.444	1	0.694

Table 4. Pairwise FST estimates from 5 microsatellite loci for populations in this study.

	Grant Park	Sedgwick Reserve	Anacapa	Santa Cruz East	Santa Cruz West	Santa Rosa	San Miguel	Santa Barbara
Grant Park	0							
Sedgwick Reserve	0.0397	0						
Anacapa	0.1675	0.14	0					
Santa Cruz East	0.1592	0.1171	0.088	0				
Santa Cruz West	0.177	0.1525	0.121	0.0081	0			
Santa Rosa	0.1099	0.084	0.1314	0.0535	0.0674	0		
San Miguel	0.1723	0.1828	0.1444	0.131	0.1141	0.1371	0	
Santa Barbara	0.1418	0.1475	0.1641	0.0959	0.1139	0.078	0.1328	0

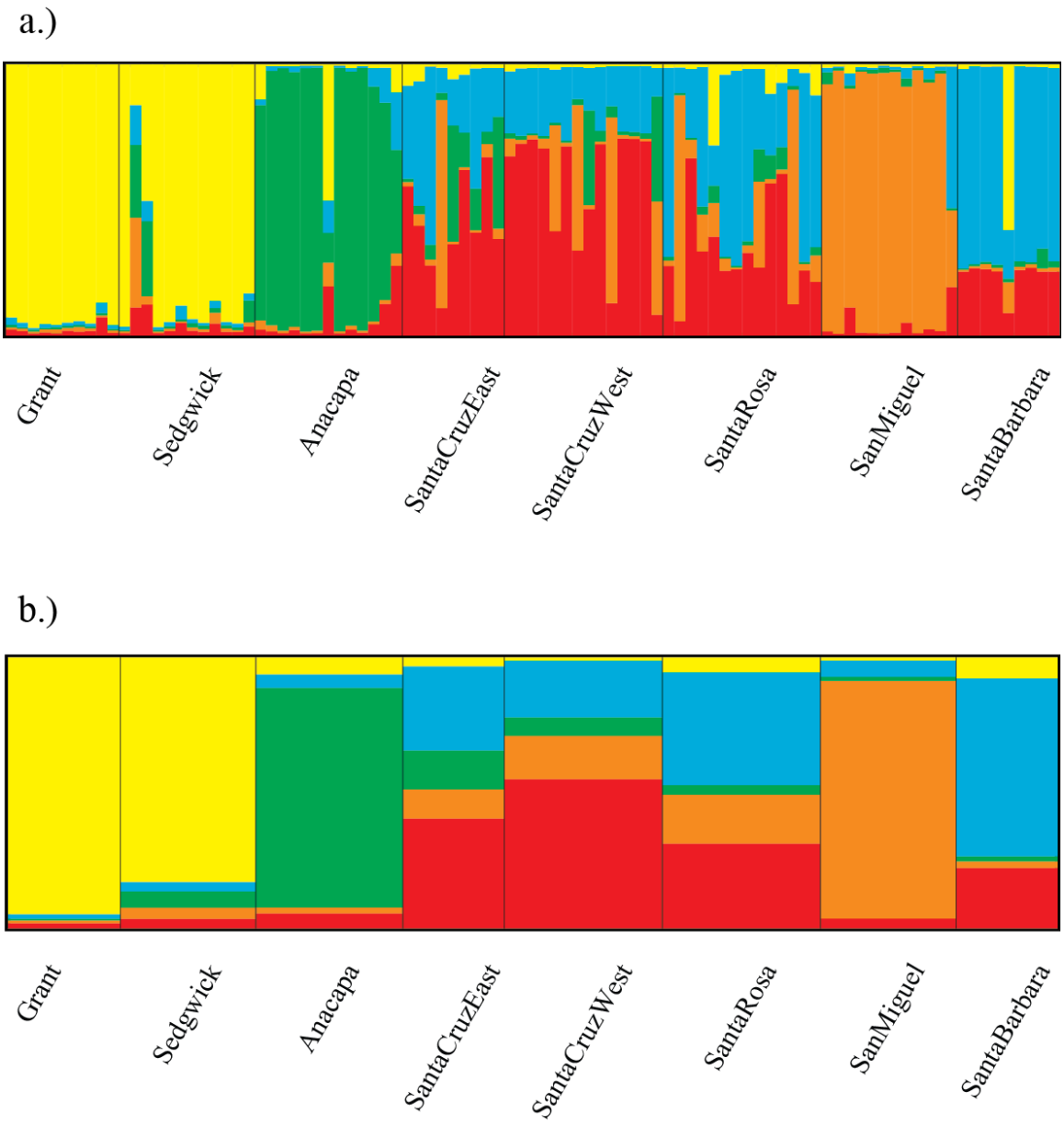


Figure 11. Inferred clusters based on a STRUCTURE analysis for K=5 clusters. Each individual is represented by a single vertical line in (a.) whereas the cluster assignments are averaged across all individuals in the population for (b.). The size of each color for an individual or population represents the probability of membership within the cluster represented by that color.

4.4 Discussion

The preliminary results from an analysis of 5 microsatellite loci are largely in agreement with the results from the previous chapter, but the microsatellite analyses provide some unique insights into both the evolutionary history and the current diversity of the northern Channel Islands. Similarly to the mitochondrial analyses, the most significant differentiation was found between island and mainland populations. The mainland had higher values for most measures of genetic diversity, pairwise F_{ST} values were higher for mainland-island comparisons than they were for most inter-island comparisons, and the STRUCTURE analysis found that the mainland and island formed distinct clusters.

In a similar vein, measures of island genetic diversity, and inter-island F_{ST} values were generally lower, and the STRUCTURE analysis found little differentiation between Santa Cruz, Santa Rosa and, to a lesser extent, Santa Barbara Island. This is in line with the mtDNA analyses, which found low levels of differentiation between different island populations. The fact that Santa Barbara Island shares cluster assignments with Santa Cruz and Santa Rosa, both of which are more similar to the northern mainland, supports a northern origin for Santa Barbara Island deer mice. That the San Miguel population is significantly different from Santa Cruz, Santa Rosa and Santa Barbara in the STRUCTURE analysis suggests that the Santa Barbara Island population may have diverged after the breakup of Santarosae, a fact that is backed up by the higher pairwise

F_{ST} value for Santa Barbara and San Miguel than for Santa Barbara and any of the other northern islands.

In the previous chapter, I found Anacapa to be significantly differentiated from the other islands. The STRUCTURE analysis in this chapter also recognized Anacapa as a unique cluster, but the other analyses in this chapter tell a different story. While the mitochondrial genetic diversity measurements were much lower for Anacapa than they were for the other islands, microsatellite diversity measurements for Anacapa were very similar to the measurements for other islands. Anacapa had an average number of private alleles, an average allelic richness, and average measurements of heterozygosity. Additionally, inter-island microsatellite F_{ST} values for Anacapa were in line with those found between other northern islands, and inter-island values between Anacapa and the other northern islands increased with physical distance, as would be expected in an isolation by distance model (something not tested in this chapter). This evidence seems to suggest that while Anacapa may have undergone a recent loss of mitochondrial diversity, there was not a similar loss of nuclear diversity, an encouraging sign for the management and conservation of Anacapa mice, a species of special concern in California (Howald et al. 2009).

4.5 Comparisons with other Channel Islands species

To better understand the phylogeography of *Peromyscus maniculatus* across the Channel Islands, it is useful to compare their phylogeography with the molecular analyses done for other endemic Channel Island species and subspecies. Plants represent the overwhelming majority of endemic species and subspecies across the

Channel Islands (cite "A checklist of vascular plants of channel islands national park" 1997), but only a few studies have examined the genetics of multi-island species. A phylogenetic study of the *Quercus* subgenus found that, like *Peromyscus*, populations of the Channel Island oak (*Quercus tomentella*) on Anacapa and Santa Cruz islands were more closely related to each other than either was to a population on San Clemente island (Manos, Doyle, and Nixon 1999; Ashley and Wills 1987). A more population-oriented study of the blue wild rye (*Elymus glaucus*) found that island-mainland differences, the main source of variation for *Peromyscus*, explained a relatively small amount of the genetic variance while most of the differentiation was found between populations (Hufford, Mazer, and Hodges 2014). Finally, an analysis of multiple species within the genus *Acmispon* found significantly lower levels of genetic differentiation within populations of the island endemic *A.dendroideus* when compared to its closest mainland relative, but no significant differences when comparing island and mainland populations of *A.argophyllus* (McGlaughlin 2014).

Several recent bird phylogeographies encompassing Channel Island populations show varying degrees of geographic differentiation largely based on the dispersal abilities of the different species. The horned lark (*Eremophila alpestris*), a long-range disperser, showed little geographic differentiation between island and mainland populations, with island populations sharing haplotypes with mainland populations as far away as Oregon and Nevada and exhibiting relatively few unique island haplotypes.

In fact, a phylogenetic analysis of horned lark populations found the Channel Island subspecies *E.alpestris insularis* to be polyphyletic, representing either multiple colonization events or incomplete lineage sorting on the islands (Mason et al. 2014). However, genetic evidence suggests that populations of horned larks have been present on the islands for over 300,000 years. On the other hand, the endemic island scrub jay (*Aphelocoma insularis*), a short-range disperser found only on Santa Cruz island, is highly diverged from its mainland counterpart, sharing no mitochondrial haplotypes with western scrub jays despite only 150,000 years of divergence between the two lineages (Delaney and Wayne 2005). Somewhere in the middle, the endemic subspecies of loggerhead shrike and song sparrow, two other short range dispersers, show patterns more in line with those seen for *Peromyscus maniculatus.* Both subspecies exhibit significant differentiation between northern and southern island populations, and they share few or no haplotypes with their nearest mainland relatives (Caballero and Ashley 2011; A. Wilson 2008).

While there are relatively few reptiles found on the Channel Islands (and therefore relatively few studies as well), a good deal of attention has been paid to the endemic island night lizard (*Xantusia riversiana*), found on San Clemente, San Nicolas and Santa Barbara islands, both on its own and in larger phylogenetic studies of the night lizard clade. An early work focusing on *X.riversiana* found significant morphological differentiation between the three island populations, and significant

genetic differentiation between island and mainland night lizards, but very little genetic differentiation between island populations based on electrophoretic analyses of proteins. Based on their knowledge of divergence times in the *Xantusia* clade, they estimated that island night lizards diverged from their mainland relatives 10-15 million years ago, while the island populations had been diverged from each other for less than a million years (Bezy et al. 1980). More recent phylogenetic work has found the populations on Santa Barbara and San Clemente islands to be more related to each other than either are to the population on San Nicolas (Noonan et al. 2013). They determined that, due to geological evidence suggesting that Santa Barbara and San Nicolas were entirely submerged at points during the Pleistocene, that San Clemente was the source population for all island night lizards, harboring a stable population since the late Miocene. The only other study focused on Channel Island reptiles examined the genetics of the side-blotched lizard (*Uta stansburiana*) and alligator lizard (*Elgaria multicarinata*) and, despite small sample sizes, they found some evidence for long-standing populations of *U.stansburiana* on San Clemente and Santa Catalina islands and *E.multicaranata* on San Nicolas island while single individuals of *U.stansburiana* on Santa Cruz and San Nicolas islands and *E.multicaranata* on San Miguel were identical to mainland individuals (Mahoney, Parks, and Fellers 2003). This is good evidence for a very modern introductions, especially since Santa Cruz and San Nicolas have been the

two islands with the most human traffic over the last fifty years and *U.stansburiana* was not even thought to be found on San Nicolas as late as the 1980's.

The phylogeographies of plants, birds and lizards across the Channel Islands provide a context for the evolution of deer mice on the islands, but the deeper evolutionary history of these taxa on the islands along with their greater dispersal abilities make it difficult to draw direct comparisons with *Peromyscus*. More important to this study are the comparisons with other terrestrial mammals found on the Channel Islands. Of the three other endemic mammals found on multiple islands, the western harvest mouse (*Reithrodontomys megalotis*) has received the least attention. In the only study of Channel Islands harvest mice, Mary Ashley (1989) used the same restriction enzymes from her work on *Peromyscus* to compare patterns of mtDNA differentiation between 5 island harvest mice from Santa Catalina and 4 mainland individuals. Unlike *Peromyscus*, however, eight of the nine individuals examined shared the exact same restriction fragment patterns. The remaining individual (an island individual) only differed in one restriction fragment, which Ashley suggested was due to a single base-pair substitution. The differences observed between deer mice and harvest mice on Santa Catalina suggest that harvest mice have a much shorter history on the Channel Islands. Their absence in the fossil record and their disjunct distribution across Santa Cruz, San Clemente and Santa Catalina islands point to a relatively recent introduction, most likely mediated by human transport from the mainland to each island individually.

In a similar vein, a recent analysis of island spotted skunks, which are now found on two islands and historically on a third, found relatively little differentiation between populations found on Santa Cruz and Santa Rosa islands (Floyd et al. 2011). In fact, the differentiation observed between the two islands was similar to the differentiation observed between the islands and the mainland, suggesting that island populations have been separated from each other as long as they have been separated from the mainland. However, unlike harvest mice, the presence of skunks on Santa Cruz and Santa Rosa islands, and their presence historically on San Miguel island does not immediately suggest multiple human introductions. A single transportation event taking place around the time of the breakup of Santarosae would be sufficient to explain the distribution of skunks across the islands.

Perhaps the most interesting phylogeographic comparison for Channel Islands deer mice is the endemic island fox (*Urocyon littoralis*). Island foxes are widespread, found on 6 of the 8 islands, and, like *Peromyscus*, they are thought to have been transported between islands by early humans (Collins 1982). In the most comprehensive molecular analysis of island foxes, Wayne et al. (1991) found them to be significantly differentiated from their closest mainland relative, the gray fox (*Urocyon cinereoargenteus*). They shared no mtDNA haplotypes with mainland gray foxes (similar to *Peromyscus*), and each island had a relatively small number of haplotypes. Unlike *Peromyscus*, however, there was relatively little geographic structure to the genetic

patterns they observed. San Miguel had a haplotype that was not seen on the other northern islands, but that was fixed on San Clemente island. Santa Rosa and Santa Cruz islands shared a haplotype with San Miguel, but they shared that haplotype, along with another haplotype not found on San Miguel, with Santa Catalina island. A more recent study analyzing 19 microsatellite loci found considerably more geographic structure, correctly identifying the origin of 181 out of 183 foxes sequenced, and showing strong support for a northern clade and a southern clade, similar to the pattern observed for *Peromyscus* (Goldstein et al. 1999). Whole mitochondrial genome sequences currently being analyzed by Courtney Hoffman and her colleagues for foxes (island and mainland) across western North America have the potential to further solidify the geographic relationships identified in previous studies, and their analysis of ancient mitochondrial genomes will give some insight into the origin and divergence of foxes across the islands (Courtney Hoffman, personal communication).

It is interesting to compare the phylogeography of these species with that of the only other mammal found on all eight of the Channel Islands, early humans. Like many other island species, humans have a much longer history on the northern islands, with pre-Clovis settlements at Arlington Springs dated at 13,000 years ago. Settlements on the southern islands are much more recent (e.g. 4,000 years ago on Santa Barbara) and they were made possible by advances in dispersal ability, namely the superior maritime

technology of the Chumash people, who replaced the pre-Clovis settlements on the islands about 11,000 years ago (Braje et al. 2010).

Although it is difficult to draw generalities from the diverse group of organisms found across the Channel Islands, two main patterns do emerge. First, as one would expect, as long as a species has a limited dispersal ability, there appears to be significant differentiation between island and mainland populations. Second, species found on both the northern and southern islands tend to be significantly differentiated between the regions. The northern islands tend to be more closely related to each other, while the relationships between the southern islands depends more on the dispersal abilities of the species in question.

4.6 Future Directions and Conclusions

While the microsatellite analyses have already yielded some intriguing results, there are still several avenues for investigation that remain unexplored. Currently, this work analyzes five microsatellite loci. Additional loci would put this study more in line with previous population-level studies of *Peromyscus*, providing finer resolution especially for the STRUCTURE analysis (Pritchard, Stephens, and Donnelly 2000). In the absence of more molecular data, principle coordinates analyses have proven useful in differentiating populations in previous studies (Yang and Kenagy 2009; Ozer, Gellerman, and Ashley 2011), and other tests (such as the previously mentioned test for isolation by distance) may also provide additional insight into the genetic structure of these populations. Additionally, while this study found significant differentiation between island and mainland populations, the role

of migration (human-mediated or otherwise) in this system remains unclear. Programs like IMa2 (Hey and Nielsen 2007) use a coalescent framework to measure effective population sizes and migration rates simultaneously while allowing the rates to change over time. IMa2 would also provide an estimate of divergence times for each of the populations, something that could have important repercussions for future studies of the Channel Islands.

While the findings in this chapter are only preliminary, they underscore the importance of a multi-locus approach to phylogeography. Several of the findings agree with the previous chapter; there are significant differences between island and mainland populations, island populations seem to share their history with the mainland to the north, and Santa Barbara Island deer mice have a more recent common ancestor with the northern island populations than they do with either mainland population. Other findings, such as the distinct clustering of San Miguel Island and the unexpected genetic diversity of Anacapa mice, show how different aspects of an organism's history can produce very different molecular signatures across nuclear and mitochondrial loci. Integrating these signals creates a broader understanding of the system, and it provides a more cohesive framework on which future research can build.

Appendix A. Reading the classification tree in Chapter One.

The classification tree (Figure 3 in main text) is meant to be read from top to bottom with the splits on the tree progressively subdividing the sample into different groups based on the values of their predictor variables. As the uppermost node of the tree indicates, the total sample of 135 records of insular rodents consisted of 78 cases (to the left of the slash) in which size was bigger on the island and 57 cases (to the right) in which the insular form was smaller. Terminal (rectangular) nodes show the results (predicted to be bigger or smaller on islands) of the classification process: for example, using "Family" as the predictor, the model identified 1+12=13 cases of island forms it would predict to be small on the basis of membership in the families Echimyidae, Gliridae, Heteromyidae, and Hystricidae. The remaining cases (77 big and 45 small) underwent further sorting using additional variables. Following the tree further down for members of the families Castoridae, Cricetidae, Muridae and Sciuridae, Mass was used to predict size reduction (to the right) in island populations whose mainland forms weigh more than 253 grams, and these 11+18 =29 cases underwent another step in which the 15 cases with 2 or more competitors present on the island (and, as before, where the mainland form weighs more than 253 grams) were predicted to be small. Of these 15 cases, 12 were predicted correctly. The rest of the tree proceeds in a similar fashion, with different variables splitting different groups of samples. Altogether, 57 + 7 + 11 + 8 + 4 + 13 + 12 = 112, or 83% of the total sample of 135 cases, were correctly predicted by this model.

Appendix B. Supplementary tables and figures for Chapter One.

Table 5. Loadings of original BioClim variables on PC1 of principal component analyses (using correlation matrix) for precipitation and temperature.

Precipitation (proportion of variance explained by PC1 = 0.931)	
Annual Precipitation	0.85915
Precipitation of Wettest Month	0.12855
Precipitation of Driest Month	0.03695
Precipitation Seasonality (Coefficient of Variation)	-0.01144
Precipitation of Wettest Quarter	0.32637
Precipitation of Driest Quarter	0.13444
Precipitation of Warmest Quarter	0.17856
Precipitation of Coldest Quarter	-0.29558
Temperature (proportion of variance explained by PC1 = 0.997)	
Annual Mean Temperature	0.02055
Mean Diurnal Range	-0.00107
Isothermality	0.06398
Temperature Seasonality	-0.99575
Max Temperature of Warmest Month	0.00689
Min Temperature of Coldest Month	0.03448
Temperature Annual Range	-0.02759
Mean Temperature of Wettest Quarter	0.01311
Mean Temperature of Driest Quarter	0.02549
Mean Temperature of Warmest Quarter	0.00848
Mean Temperature of Coldest Quarter	0.03287

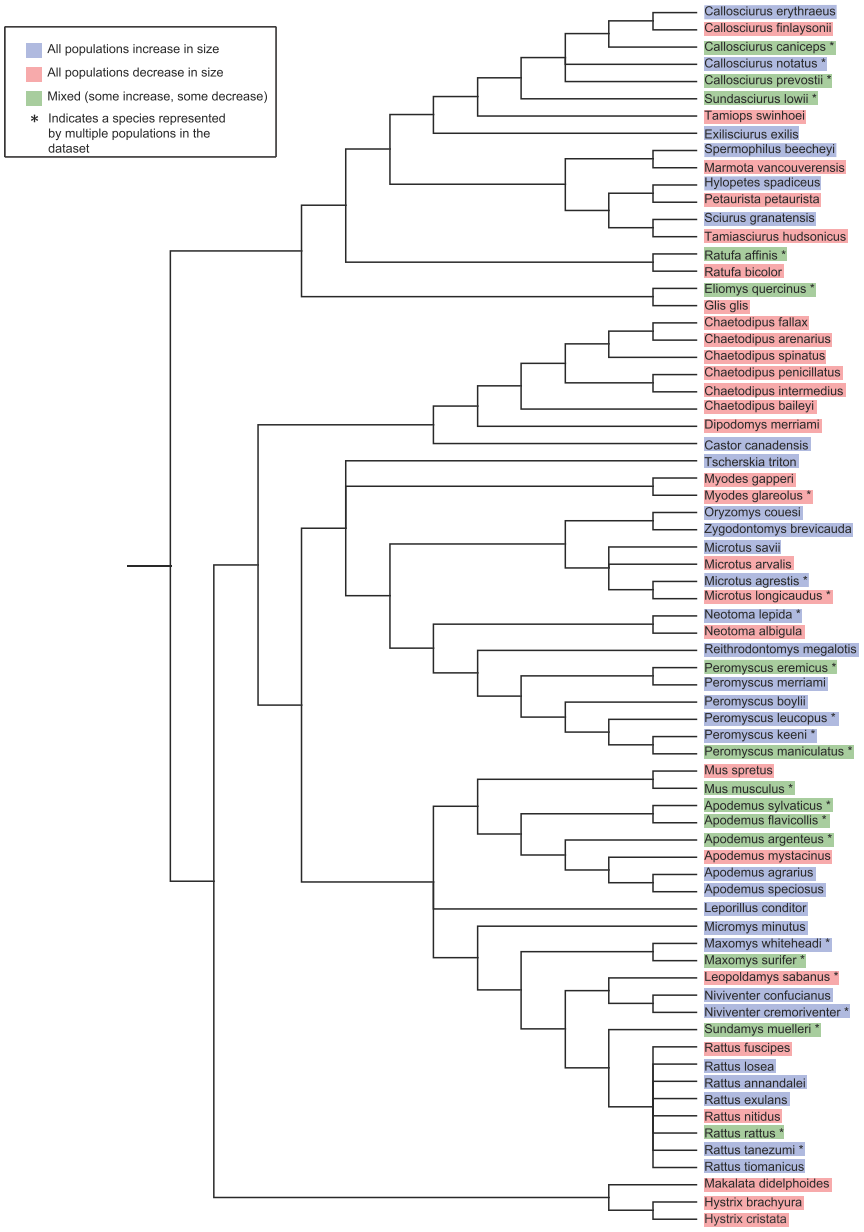
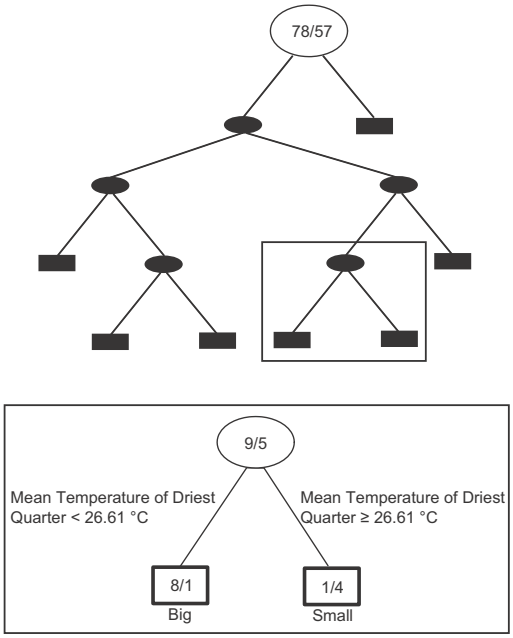
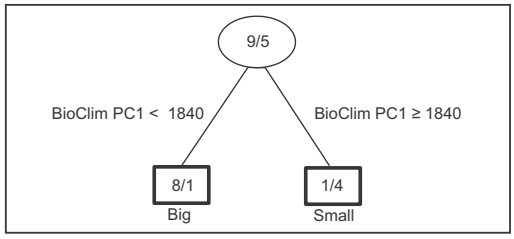


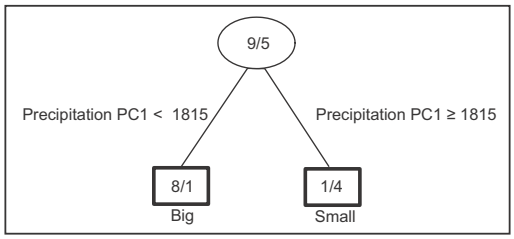
Figure 12. A phylogenetic tree of all rodent species used in this study. Colors (blue, red, green) indicate the direction of size change for each species (all populations increase, all decrease, or mixed increase and decrease). All 8 families from Meiri's dataset (2008) are represented in this subsample. We should emphasize that the presence or absence of phylogenetic structuring is an attribute of a particular dataset, is subject to sampling bias, and using different or additional exemplars of the same higher taxa could yield a different conclusion. The results of our within-island and within-species tests are presumably less subject to such bias.



a. Analysis with the 19 original BioClim variables



b. Analysis with a single PCA for all BioClim variables



c. Analysis with separate PCAs for temperature and precipitation variables

Figure 13. Stability of classification tree using different climate variables. We ran classification-tree analyses using several different sets of climatic variables (the original 19 BioClim variables, PC1 from a single PCA using all 19 variables, and PC1 from separate PCAs for temperature and precipitation variables). The structure of the tree did not change across analyses. The only difference between the trees was the variable used at the node illustrated above.

Appendix C. Assessing the phylogenetic structure of size change in Chapter Two

We performed several permutation tests to answer three distinct questions about the structure of our dataset: Is there any phylogenetic structure to the tendency to undergo extreme size change? Is there a tendency for a single pattern (extreme or normal insular size) to predominate within a species? Among those species exhibiting extreme insular size, do multiple populations of the same species tend to undergo extreme size change in the same direction (big or small)?

To answer question 1, we assembled a phylogenetic tree by splicing together phylogenies from multiple sources (Figure 14, see chapter one for references). Each species was represented by a single terminal branch and each terminal branch was assigned a character state based on whether at least one of its populations undergoes extreme size change. We calculated the total length of the tree in character transition steps and permuted the tips 1,000 times to compare a distribution of tree lengths with the length of the original tree (see chapter two for details on the methods). Phylogenetic structuring of extreme size would be indicated by an original tree with significantly fewer steps than expected by chance, but the original tree fell well within the distribution produced by random tree permutations ($p=0.48$), indicating that any given clade is no more likely to undergo extreme size change than any other clade.

For question two, we created a tree diagram consisting of a basal polytomy of 67 branches (for the 67 species in our dataset). An additional polytomy was added to each branch, corresponding to the number of populations for that species in our dataset. We assigned character states to each of the populations (extreme or normal), and again calculated the length of the tree in character steps and compared that length to a distribution of lengths for 1,000 trees with randomly permuted tips. A significantly smaller length for the original tree would indicate a tendency for certain species to undergo extreme size change, but the original tree again fell well within the distribution produced by random tip permuations ($p=0.42$), indicating that there is no significant trend for individual species to undergo size change that is consistently extreme.

Finally, to address question 3, we created a tree diagram similar to that for question 2, but with a basal polytomy of just 25 branches representing each of the species with multiple extreme sized populations. Again each species branched into an additional polytomy with each branch representing an extreme sized population, and character states were assigned to each population based on whether they were extremely big or extremely small, and we compared the length of the original tree to the distribution of lengths for 1,000 trees with randomly permuted tips. A significantly smaller length for the original tree would indicate a tendency for populations within a species to tend towards extreme size change in a single direction (big or small), and we did observe such a trend ($p<<0.001$), indicating that although there is no tendency for

species to undergo extreme size change, when more than one population within a species does so the tendency is to change in the same direction (big or small) every time.

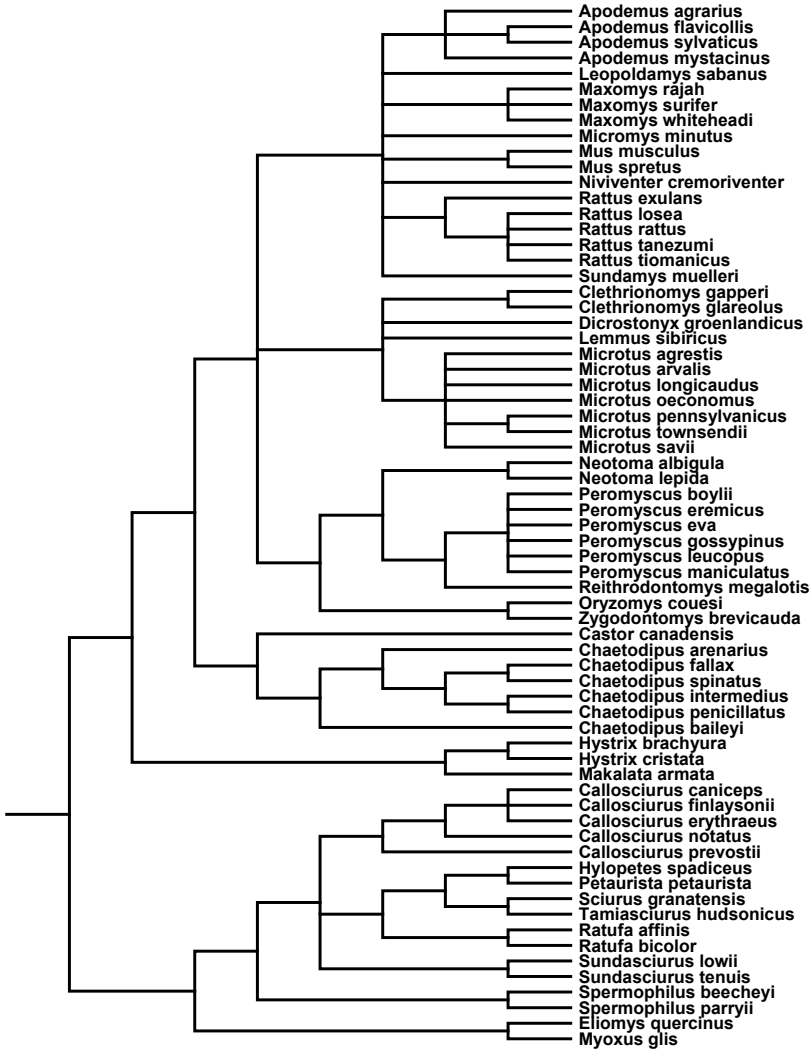


Figure 14. Phylogenetic relationships among species examined in chapter two based on the phylogenetic tree depicted in Figure 12.

Appendix D. Examining the asymmetric boundary for 'Extreme' size change in Chapter Two

Our analysis of mainland rodent size variation produced an asymmetric distribution of ratios for rodent populations compared to their mainland species average. The upper 97.5% quantile consisted of populations at least 9% bigger than their mainland species average, while the lower 2.5% quantile consisted of populations more than 20% smaller than their mainland species averages. Since there is such a discrepancy in the degree of size change included in the upper and lower quantiles, we took our original sample of 306 island species, and we examined how well both the classification tree and random forest methods performed with our asymmetric boundaries ('mainland boundaries') when compared to symmetric boundaries, which used either the upper quantile ('inclusive boundaries', considering more populations to be extreme by defining 'normal' size ratios as 0.91-1.09) or the lower quantile ('stricter boundaries', defining 'normal' size ratios as 0.80-1.20) to determine extreme size. We used the same variables as in our original analysis, with the only difference being the number of cases included in the analyses. Once the trees and random forests were built, we ran the independent data set (new data obtained from Lomolino et al. 2013) through each model to determine the predictive accuracy for each sample size. Additionally, we built 1,000 random forests from a random sampling, for each sample size, of the original 306 species to examine the impact sample size had on predictive accuracy.

For most classification tree and random forest analyses, models using the boundaries based on quantiles of the mainland body size distribution performed better than either the inclusive or the strict symmetrical boundaries (see table below). Random forests based on random (including both 'normal' and extreme) samples of the insular rodent populations were relatively similar regardless of sample size, and performed two to three times worse than any other model. This demonstrates that clearer patterns emerge when the focus is on size variation that extends beyond that commonly found among populations on the mainland.

Applying strict boundaries, as compared to the mainland quantile boundaries, is a matter of reducing the number of Big island populations in the sample. In general, doing so raised the misclassification rate, which is to be expected because the trees in these analyses were found to be more successful in predicting size increase than decrease on islands. Predictive classification of the independent data appears to have improved, but at this small sample size the difference between the rate of misclassification using strict boundaries (~11%) and that produced using mainland quantile (~15%) amounts to a single case.

When more inclusive boundaries are applied, the large disparity in predictive accuracy between classification trees and random forests suggests that classification trees produced using these boundaries are relatively unstable. Using inclusive boundaries adds Small island populations to the sample, and the likelihood that

undetected subadult or juvenile individuals have been included and influenced population averages is greater in Small island populations whose Size Ratios fall closer to 1.

Table 6. Effects of Shifting the "Extreme" Boundaries

# Large/# Small	Strict Boundaries (n=120, new*=27) (83/37, 22/5)	Mainland quantile Boundaries (n=169, new*=39) (132/37, 34/5)	Inclusive Boundaries (n=206, new*=43) (132/74, 34/9)
Classification Tree Misclassification Rate	15.00%	10.00%	16.02%
Random Forest Misclassification Rate	16.38%	13.11%	26.28%
Predictive Misclassification Rate for Independent Data	11.11%	15.38%	23.26%
Random Forest Misclassification Rate for Independent Data	15.67%	11.30%	18.79%
Random Sample Mean Misclassification Rate (Random Forest)	35.67%	34.37%	33.77%
	*'new' refers to independent data set		

Appendix E. Supplementary table for Chapter Two

Table 7. Loadings of island terrain heterogeneity variables on PC1 of principal component analysis (using correlation matrix) for Chapter Two.

Island Terrain Heterogeneity (proportion of variance explained by PC1 = 0.766)	
Island Area	0.598
Maximum Elevation	0.450
Mean Elevation	0.379
Median Elevation	0.391
Standard Deviation of Elevation	0.379

References

Adams, A. Leith. 1874. "I. On the Dentition and Osteology of the Maltese Fossil Elephants, Being a Description of Remains Discovered by the Author in Malta, between the Years 1860 and 1866." *The Transactions of the Zoological Society of London* 9 (1): 1–124. doi:10.1111/j.1096-3642.1874.tb00235.x.

Adler, G.H. GH, and Richard Levins. 1994. "The Island Syndrome in Rodent Populations." *Quarterly Review of Biology* 69 (4). JSTOR: 473–90.

Alexander, Lois F., and Brett R. Riddle. 2005. "Phylogenetics of the New World Rodent Family Heteromyidae." *Journal of Mammalogy* 86 (2): 366–79. doi:10.1644/BER-120.1.

Ashley, Mary V. 1989. "Absence of Differentiation in Mitochondrial DNA of Island and Mainland Harvest Mice , Reithrodontomys Megalotis." *Journal of Mammalogy* 70 (2): 383–86.

Ashley, Mary V, and Christopher Wills. 1987. "Analysis of Mitochondrial DNA Polymorphisms Among Channel Island Deer Mice." *Evolution* 41 (4): 854. doi:10.2307/2408893.

Avise, J.C. 2000. *Phylogeography: The History and Formation of Species*. Harvard Univ Pr.

Ballard, J. William O, and Michael C. Whitlock. 2004. "The Incomplete Natural History of Mitochondria." *Molecular Ecology*. doi:10.1046/j.1365-294X.2003.02063.x.

Barraclough, T. G., a. P. Vogler, and P. H. Harvey. 1998. "Revealing the Factors That Promote Speciation." *Philosophical Transactions of the Royal Society B: Biological Sciences* 353 (1366): 241–49. doi:10.1098/rstb.1998.0206.

Benton, Michael J., Zoltan Csiki, Dan Grigorescu, Ragna Redelstorff, P. Martin Sander, Koen Stein, and David B. Weishampel. 2010. "Dinosaurs and the Island Rule: The Dwarfed Dinosaurs from Haţeg Island." *Palaeogeography, Palaeoclimatology, Palaeoecology* 293 (3-4). Elsevier B.V. 438–54. doi:10.1016/j.palaeo.2010.01.026.

Bezy, R L, G C Gorman, G A Adest, and Y J Kim. 1980. "Divergence in the Island Night Lizard Xantusia Riversiana (Sauria: Xantusiidae)." In *The California Islands, Proceedings of a Multidisciplinary Symposium*, edited by Dennis M Power, 565–83. Santa Barbara, CA: Santa Barbara Museum of Natural History.

Blanga-Kanfi, Shani, Hector Miranda, Osnat Penn, Tal Pupko, Ronald W DeBry, and Dorothée Huchon. 2009. "Rodent Phylogeny Revised: Analysis of Six Nuclear Genes from All Major Rodent Clades." *BMC Evolutionary Biology* 9 (1). BioMed Central: 71. doi:10.1186/1471-2148-9-71.

Boyer, Alison G. 2010. "Consistent Ecological Selectivity through Time in Pacific Island Avian Extinctions." *Conservation Biology : The Journal of the Society for Conservation Biology* 24 (2): 511–19. doi:10.1111/j.1523-1739.2009.01341.x.

Bradley, Robert D, Nevin D Durish, Duke S Rogers, Jacqueline R Miller, Mark D Engstrom, and C William Kilpatrick. 2007. "Toward a Molecular Phylogeny for Peromyscus: Evidence from Mitochondrial Cytochrome-B Sequences." *Journal of Mammalogy* 88 (5): 1146–59. doi:10.1644/06-MAMM-A-342R.1.

Braje, Todd J, Julia G Costello, Jon M Erlandson, Michael A Glassow, John R Johnson, Don P Morris, Jennifer E Perry, and Torben C Rick. 2010. *Channel Islands National Park Archaeological Overview and Assessment*. Ventura, CA.

Britton-Davidian, J, J Catalan, M da Graça Ramalhinho, G Ganem, J C Auffray, R Capela, M Biscoito, J B Searle, and M da Luz Mathias. 2000. "Rapid Chromosomal Evolution in Island Mice." *Nature* 403: 158. doi:10.1038/35003116.

Bromham, Lindell, and Marcel Cardillo. 2007. "Primates Follow the 'Island Rule': Implications for Interpreting Homo Floresiensis." *Biology Letters* 3 (4): 398–400. doi:10.1098/rsbl.2007.0113.

Brown, James H., and D W Davidson. 1977. "Competition between Seed-Eating Rodents and Ants in Desert Ecosystems." *Science* 196 (4292): 880–82. doi:10.1126/science.196.4292.880.

Busk, George. 1868. "Description of the Remains of Three Extinct Species of Elephant, Collected by Capt. Spratt, C.B., R.N., in the Ossiferous Cavern of Zebbug, in the Island of Malta." *The Transactions of the Zoological Society of London* 6 (5): 227–306. doi:10.1111/j.1096-3642.1868.tb00578.x.

Caballero, Isabel C., and Mary V. Ashley. 2011. "Genetic Analysis of the Endemic Island Loggerhead Shrike, Lanius Ludovicianus Anthonyi." *Conservation Genetics* 12 (6): 1485–93. doi:10.1007/s10592-011-0247-4.

CEPF Report: A New Model for Global Conservation. 2011.

Clement, M, D Posada, and K A Crandall. 2000. "TCS: A Computer Program to Estimate Gene Genealogies." *Molecular Ecology* 9 (10): 1657–59.

Collins, Paul William. 1982. "Origin and Differentiation of the Island Fox: A Study of Evolution in Insular Populations". University of California, Santa Barbara.

Conroy, Chris J, and Joseph A Cook. 2000. "Molecular Systematics of a Holarctic Rodent (Microtus : Muridae)." *Journal of Mammalogy* 81: 344–59. doi:10.1644/1545-1542(2000)081<0344:MSOAHR>2.0.CO;2.

Cutler, D. Richard, Thomas C. Edwards, Karen H. Beard, Adele Cutler, T Kyle, Jacob Gibson, Joshua J. Lawler, H Beard, T Hess, and Kyle T. Hess. 2007. "Random Forests for Classification in Ecology." *Ecology* 88 (11): 2783–92. doi:10.1890/07-0539.1.

Darwin, Charles. 1859. *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life. Darwin*. Vol. 5.

Davidson, Ana D, Marcus J Hamilton, Alison G Boyer, James H. Brown, and Gerardo Ceballos. 2009. "Multiple Ecological Pathways to Extinction in Mammals." *Proceedings of the National Academy of Sciences of the United States of America* 106 (26): 10702–5. doi:10.1073/pnas.0901956106.

Delaney, Kathleen S., and Robert K. Wayne. 2005. "Adaptive Units for Conservation: Population Distinction and Historic Extinctions in the Island Scrub-Jay." *Conservation Biology* 19 (2): 523–33. doi:10.1111/j.1523-1739.2005.00424.x.

Dragoo, Jerry W., J Alden Lackey, Kathryn E Moore, Enrique P Lessa, Joseph a Cook, and Terry L Yates. 2006. "Phylogeography of the Deer Mouse (Peromyscus Maniculatus) Provides a Predictive Framework for Research on Hantaviruses." *The Journal of General Virology* 87 (Pt 7): 1997–2003. doi:10.1099/vir.0.81576-0.

Drummond, Alexei J, and Andrew Rambaut. 2007. "BEAST: Bayesian Evolutionary Analysis by Sampling Trees." *BMC Evolutionary Biology* 7 (January): 214. doi:10.1186/1471-2148-7-214.

Drummond, Alexei J., Marc A. Suchard, Dong Xie, and Andrew Rambaut. 2012. "Bayesian Phylogenetics with BEAUti and the BEAST 1.7." *Molecular Biology and Evolution* 29: 1969–73. doi:10.1093/molbev/mss075.

Durst, Paul A P, and V Louise Roth. 2012. "Classification Tree Methods Provide a Multifactorial Approach to Predicting Insular Body Size Evolution in Rodents." *The American Naturalist* 179 (4): 545–53. doi:10.1086/664611.

Earl, Dent A., and Bridgett M. vonHoldt. 2012. "STRUCTURE HARVESTER: A Website and Program for Visualizing STRUCTURE Output and Implementing the Evanno Method." *Conservation Genetics Resources* 4: 359–61. doi:10.1007/s12686-011-9548-7.

Erlandson, Jon M, Torben C Rick, Todd J Braje, Molly Casperson, Brendan Culleton, Brian Fulfrost, Tracy Garcia, et al. 2011. "Paleoindian Seafaring, Maritime Technologies, and Coastal Foraging on California's Channel Islands." *Science* 331 (6021): 1181–85. doi:10.1126/science.1201477.

Evanno, G, S Regnaut, and J Goudet. 2005. "Detecting the Number of Clusters of Individuals Using the Software STRUCTURE: A Simulation Study." *Molecular Ecology* 14 (8): 2611–20. doi:10.1111/j.1365-294X.2005.02553.x.

Excoffier, L, and H Lischer. 2005. "Arlequin Ver. 3.0: An Integrated Software Package for Population Genetics Data Analysis." *Evolutionary Bioinformatics Online* 1 (1). Evolutionary Bioinformatics Online: 47–50.

Floyd, Chris H., Dirk H. Van Vuren, Kevin R. Crooks, Krista L. Jones, David K. Garcelon, Natalia M. Belfiore, Jerry W. Dragoo, and Bernie May. 2011. "Genetic Differentiation of Island Spotted Skunks, Spilogale Gracilis Amphiala." *Journal of Mammalogy* 92 (1): 148–58. doi:10.1644/09-MAMM-A-204.1.

Foster, J Bristol. 1964. "Evolution of Mammals on Islands." *Nature* 202 (4929). Nature Publishing Group: 234–35. doi:10.1038/202234a0.

Fritz, Susanne A, Olaf R P Bininda-Emonds, and Andy Purvis. 2009. "Geographical Variation in Predictors of Mammalian Extinction Risk: Big Is Bad, but Only in the Tropics." *Ecology Letters* 12 (6). WILEY-BLACKWELL PUBLISHING: 538–49.

Garth, John S. 1965. "Introduction to Insular Zoology." In *1st Symposium on the Biology of the California Islands*, 181–83. National Park Service.

Gaston, Kevin J, and Tim M Blackburn. 1996. "Range Size-Body Size Relationships: Evidence of Scale Dependence." *Oikos* 75 (3). JSTOR: 479. doi:10.2307/3545889.

Gill, Ayesha E. 1976. "Genetic Divergence of Insular Populations of Deer Mice." *Biochemical Genetics* 14 (9-10): 835–48.

Goldstein, D B, G W Roemer, D a Smith, D E Reich, a Bergman, and R K Wayne. 1999. "The Use of Microsatellite Variation to Infer Population Structure and Demographic History in a Natural Model System." *Genetics* 151 (2): 797–801.

Goudet, J. 1995. "FSTAT (Version 1.2): A Computer Program to Calculate F-Statistics." *Journal of Heredity* 86: 485–86. doi:10.1371/journal.pone.0026694.

Guthrie, Daniel A. 1993. "New Information on the Prehistoric Fauna of San Miguel Island, California." In *Third California Islands Symposium: Recent Advances in Research on the California Islands*, 405–16.

Heaney, L R, E K Walker, B R Tabaranza Jr, and N R Ingle. 2002. "Mammalian Diversity in the Philippines: An Assessment of the Adequacy of Current Data." *Sylvatrop* 10 (2000): 6–27.

Heaney, Lawrence R. 1978. "Island Area and Body Size of Insular Mammals: Evidence from the Tri-Colored Squirrel (Callosciurus Prevosti) of Southeast Asia." *Evolution* 32 (1): 29–44. doi:10.2307/2407408.

Hey, Jody, and Rasmus Nielsen. 2007. "Integration within the Felsenstein Equation for Improved Markov Chain Monte Carlo Methods in Population Genetics." *Proceedings of the National Academy of Sciences of the United States of America* 104 (8). National Acad Sciences: 2785–90.

Hijmans, Robert J., Susan E. Cameron, Juan L. Parra, Peter G. Jones, and Andy Jarvis. 2005. "Very High Resolution Interpolated Climate Surfaces for Global Land Areas." *International Journal of Climatology* 25 (15): 1965–78. doi:10.1002/joc.1276.

Howald, Gregg, C. Josh Donlan, Kate R. Faulkner, Steve Ortega, Holly Gellerman, Donald a. Croll, and Bernie R. Tershy. 2009. "Eradication of Black Rats Rattus Rattus from Anacapa Island." *Oryx* 44 (01): 30. doi:10.1017/S003060530999024X.

Hufford, Kristina M, Susan J Mazer, and Scott a Hodges. 2014. "Genetic Variation among Mainland and Island Populations of a Native Perennial Grass Used in Restoration." *AoB Plants* 6 (January): 1–12. doi:10.1093/aobpla/plt055.

Jaarola, M., N. Martínková, I. Gündüz, C. Brunhoff, J. Zima, A. Nadachowski, G. Amori, et al. 2004. "Molecular Phylogeny of the Speciose Vole Genus Microtus(Arvicolinae, Rodentia) Inferred from Mitochondrial DNA Sequences." *Molecular Phylogenetics and Evolution* 33: 647–63. doi:10.1016/j.ympev.2004.07.015.

Jakobsson, Mattias, and Noah A. Rosenberg. 2007. "CLUMPP: A Cluster Matching and Permutation Program for Dealing with Label Switching and Multimodality in Analysis of Population Structure." *Bioinformatics* 23: 1801–6. doi:10.1093/bioinformatics/btm233.

Jansa, Sharon A, and Marcelo Weksler. 2004. "Phylogeny of Muroid Rodents: Relationships within and among Major Lineages as Determined by IRBP Gene Sequences." *Molecular Phylogenetics and Evolution* 31 (1): 256–76. doi:10.1016/j.ympev.2003.07.002.

Jombart, Thibaut. 2008. "Adegenet: A R Package for the Multivariate Analysis of Genetic Markers." *Bioinformatics* 24: 1403–5. doi:10.1093/bioinformatics/btn129.

Jones, Martin J., Alan Fielding, and Matthew Sullivan. 2006. "Analysing Extinction Risk in Parrots Using Decision Trees." *Biodiversity and Conservation* 15 (6): 1993–2007. doi:10.1007/s10531-005-4316-1.

King, John Arthur, ed. 1968. *Biology of Peromyscus (Rodentia)*. *Society*. Stillwater, OK: American Society of Mammalogists. doi:http://dx.doi.org/10.5962/bhl.title.39510.

Landry, S.O. O. 1970. "The Rodentia as Omnivores." *Quarterly Review of Biology* 45 (4). JSTOR: 351–72.

Larkin, M. A., G. Blackshields, N. P. Brown, R. Chenna, P. A. Mcgettigan, H. McWilliam, F. Valentin, et al. 2007. "Clustal W and Clustal X Version 2.0." *Bioinformatics* 23: 2947–48. doi:10.1093/bioinformatics/btm404.

Lawlor, TE. 1982. "The Evolution of Body Size in Mammals : Evidence from Insular Populations in Mexico." *American Naturalist* 119 (1): 54–72.

Liaw, Andy, and Matthew Wiener. 2002. "Classification and Regression by randomForest." *R News* 2 (3): 18–22.

Lomolino, MV. 1985. "Body Size of Mammals on Islands: The Island Rule Reexamined." *American Naturalist* 125 (2): 310–16.

Lomolino, Mark V. 1985. "Body Size of Mammals on Islands: The Island Rule Reexamined." *The American Naturalist* 125 (2). JSTOR: 310–16. doi:10.1086/282871.

———. 2005. "Body Size Evolution in Insular Vertebrates: Generality of the Island Rule." *Journal of Biogeography* 32 (10): 1683–99. doi:10.1111/j.1365-2699.2005.01314.x.

Lomolino, Mark V, Dov F. Sax, Maria Rita Palombo, and Alexandra A Van Der Geer. 2012. "Of Mice and Mammoths: Evaluations of Causal Explanations for Body Size Evolution in Insular Mammals." *Journal of Biogeography* 39 (5): 842–54. doi:10.1111/j.1365-2699.2011.02656.x.

Lomolino, Mark V., Alexandra A. van der Geer, George A. Lyras, Maria Rita Palombo, Dov F. Sax, and Roberto Rozzi. 2013. "Of Mice and Mammoths: Generality and Antiquity of the Island Rule." Edited by Kostas Triantis. *Journal of Biogeography* 40 (8): 1427–39. doi:10.1111/jbi.12096.

Maddison, W.P., and D.R. Maddison. 2011. "Mesquite: A Modular System for Evolutionary Analysis. Version 2.75." *Http://mesquiteproject. Org*.

Mahoney, Meredith J., Duncan S. M. Parks, and Gary M. Fellers. 2003. "Uta Stansburiana and Elgaria Multicarinata on the California Channel Islands: Natural Dispersal or Artificial Introduction?" *Journal of Herpetology*. doi:10.1670/24-01A.

Manos, P S, J J Doyle, and K C Nixon. 1999. "Phylogeny, Biogeography, and Processes of Molecular Differentiation in Quercus Subgenus Quercus (Fagaceae)." *Molecular Phylogenetics and Evolution* 12 (3): 333–49. doi:10.1006/mpev.1999.0614.

Mason, Nicholas a., Pascal O. Title, Carla Cicero, Kevin J Burns, and Rauri C. K. Bowie. 2014. "Genetic Variation among Western Populations of the Horned Lark (Eremophila Alpestris) Indicates Recent Colonization of the Channel Islands off Southern California, Mainland-Bound Dispersal, and Postglacial Range Shifts." *The Auk* 131 (2): 162–74. doi:10.1642/AUK-13-181.1.

Matisoo-Smith, E, R M Roberts, G J Irwin, J S Allen, D Penny, and D M Lambert. 1998. "Patterns of Prehistoric Human Mobility in Polynesia Indicated by mtDNA from the Pacific Rat." *Proceedings of the National Academy of Sciences of the United States of America* 95: 15145–50. doi:10.1073/pnas.95.25.15145.

McClain, Craig R, Paul A P Durst, Alison G Boyer, and Clinton D Francis. 2013. "Unravelling the Determinants of Insular Body Size Shifts." *Biology Letters* 9 (1): 20120989. doi:10.1098/rsbl.2012.0989.

McGlaughlin, ME. 2014. "Do the Island Biogeography Predictions of MacArthur and Wilson Hold When Examining Genetic Diversity on the near Mainland California Channel Islands? Examples." *Botanical Journal of …*, 289–304.

McNab, Brian K. 1963. “Bioenergetics and the Determination of Home Range Size.” *The American Naturalist*. doi:10.1086/282264.

Meiri, Shai, Natalie Cooper, and Andy Purvis. 2008. “The Island Rule: Made to Be Broken?” *Proceedings. Biological Sciences / The Royal Society* 275 (1631). The Royal Society: 141–48. doi:10.1098/rspb.2007.1056.

Meiri, Shai, Tamar Dayan, and Daniel Simberloff. 2006. “The Generality of the Island Rule Reexamined.” *Journal of Biogeography* 33 (9): 1571–77. doi:10.1111/j.1365-2699.2006.01523.x.

Mercer, John M, and V. Louise Roth. 2003. “The Effects of Cenozoic Global Change on Squirrel Phylogeny.” *Science (New York, N.Y.)* 299 (5612): 1568–72. doi:10.1126/science.1079705.

Michaux, Jacques, Pascale Chevret, and Sabrina Renaud. 2007. “Morphological Diversity of Old World Rats and Mice (Rodentia, Muridae) Mandible in Relation with Phylogeny and Adaptation.” *Journal of Zoological Systematics and Evolutionary Research* 45 (3). Blackwell Publishing: 263–79. doi:10.1111/j.1439-0469.2006.00390.x.

Michaux, Johan R., Joëlle Goüy de Bellocq, Maurizio Sarà, and Serge Morand. 2002. “Body Size Increase in Insular Rodent Populations: A Role for Predators?” *Global Ecology and Biogeography* 11 (5): 427–36. doi:10.1046/j.1466-822x.2002.00301.x.

Millien, Virginie. 2006. “Morphological Evolution Is Accelerated among Island Mammals.” *PLoS Biology* 4 (10): e321. doi:10.1371/journal.pbio.0040321.

———. 2011. “Mammals Evolve Faster on Smaller Islands.” *Evolution; International Journal of Organic Evolution* 65 (7): 1935–44. doi:10.1111/j.1558-5646.2011.01268.x.

Millien, Virginie, and John Damuth. 2004. “Climate Change and Size Evolution in an Island Rodent Species: New Perspectives on the Island Rule.” *Evolution* 58 (6): 1353–60. doi:10.1554/03-727.

Mullen, Lynne M., Rachel J. Hirschmann, Kelly L. Prince, Travis C. Glenn, Michael J. Dewey, and Hopi E. Hoekstra. 2006. “Sixty Polymorphic Microsatellite Markers for the Oldfield Mouse Developed in Peromyscus Polionotus and Peromyscus Maniculatus.” *Molecular Ecology Notes* 6 (1): 36–40. doi:10.1111/j.1471-8286.2005.01128.x.

Nichols, Richard. 2001. "Gene Trees and Species Trees Are Not the Same." *Trends in Ecology & Evolution*. doi:10.1016/S0169-5347(01)02203-0.

Noonan, Brice P., Jennifer B. Pramuk, Robert L. Bezy, Elizabeth A. Sinclair, Kevin de Queiroz, and Jack W. Sites. 2013. "Phylogenetic Relationships within the Lizard Clade Xantusiidae: Using Trees and Divergence Times to Address Evolutionary Questions at Multiple Levels." *Molecular Phylogenetics and Evolution* 69: 109–22. doi:10.1016/j.ympev.2013.05.017.

Nylander, J.A.A. 2004. "MrModeltest v2." *Evolutionary Biology Centre, Uppsala University*.

Orrock, John L. 2010. "When the Ghost of Predation Has Passed: Do Rodents from Islands with and without Fox Predators Exhibit Aversion to Fox Cues?" *Ethology* 116 (4): 338–45. doi:10.1111/j.1439-0310.2010.01740.x.

Oshida, Tatsuo. 2004. "Phylogenetic Position of the Small Kashmir Flying Squirrel," *Canadian Journal of Zoology* 82 (8): 1336–42. doi:10.1139/z04-108.

Ozer, Fusun, Holly Gellerman, and Mary V Ashley. 2011. "Genetic Impacts of Anacapa Deer Mice Reintroductions Following Rat Eradication." *Molecular Ecology* 20 (17): 3525–39. doi:10.1111/j.1365-294X.2011.05165.x.

Palmer, Miquel. 2002. "Testing the 'Island Rule' for a Tenebrionid Beetle (Coleoptera, Tenebrionidae)." *Acta Oecologica* 23 (2): 103–7. doi:10.1016/S1146-609X(02)01140-2.

Paquette, Sebastien Rioux. 2012. "PopGenKit: Useful Functions for (batch) File Conversion and Data Resampling in Microsatellite Datasets."

Paradis, Emmanuel. 2010. "Pegas: An R Package for Population Genetics with an Integrated-Modular Approach." *Bioinformatics*. doi:10.1093/bioinformatics/btp696.

Pergams, Oliver R. W., and Mary V Ashley. 1999. "Rapid Morphological Change in Channel Island Deer Mice." *Evolution* 53 (5). JSTOR: 1573. doi:10.2307/2640902.

Pergams, Oliver R. W., Robert C. Lacy, and Mary V Ashley. 2000. "Conservation and Management of Anacapa Island Deer Mice." *Conservation Biology* 14 (3): 819–32. doi:10.1046/j.1523-1739.2000.98524.x.

Peters, Robert Henry. 1983. *The Ecological Implications of Body Size*. Cambridge University Press. doi:http://dx.doi.org/10.1017/CBO9780511608551.

Prasad, Anantha M., Louis R. Iverson, and Andy Liaw. 2006. "Newer Classification and Regression Tree Techniques: Bagging and Random Forests for Ecological Prediction." *Ecosystems* 9 (2): 181–99. doi:10.1007/s10021-005-0054-1.

Price, Samantha A, Samantha S B Hopkins, Kathleen K Smith, and V Louise Roth. 2012. "Tempo of Trophic Evolution and Its Impact on Mammalian Diversification." *Proceedings of the National Academy of Sciences* 109 (18). National Acad Sciences: 7008–12.

Pritchard, J K, M Stephens, and P Donnelly. 2000. "Inference of Population Structure Using Multilocus Genotype Data." *Genetics* 155 (2). Genetics Soc America: 945–59.

R Development Core Team. 2013. "R: A Language and Environment for Statistical Computing." *R Foundation Statistical Computing*.

Raia, Pasquale, C Barbera, and M Conte. 2003. "The Fast Life of a Dwarfed Giant." *Evolutionary Ecology* 23 (6): 867–78. doi:10.1023/A:1025577414005.

Raia, Pasquale, Francesco Carotenuto, and Shai Meiri. 2010. "One Size Does Not Fit All: No Evidence for an Optimal Body Size on Islands." *Global Ecology and Biogeography* 19 (May): 475–84. doi:10.1111/j.1466-8238.2010.00531.x.

Raia, Pasquale, and Shai Meiri. 2006. "The Island Rule in Large Mammals: Paleontology Meets Ecology." *Evolution; International Journal of Organic Evolution* 60 (8). Wiley Online Library: 1731–42.

Rice, William R. 1989. "Analyzing Tables of Statistical Tests." *Evolution* 43: 223–225. doi:10.2307/2409177.

Rick, Torben C., Jon M. Erlandson, René L. Vellanoweth, Todd J. Braje, Paul W. Collins, Daniel A Guthrie, and Thomas W. Stafford Jr. 2009. "Origins and Antiquity of the Island Fox (Urocyon Littoralis) on California's Channel Islands." *Quaternary Research* 71 (2). Elsevier B.V. 93–98. doi:10.1016/j.yqres.2008.12.003.

Rosenberg, Noah A. 2004. "DISTRUCT: A Program for the Graphical Display of Population Structure." *Molecular Ecology Notes* 4: 137–38. doi:10.1046/j.1471-8286.2003.00566.x.

Roth, V L, and W D Dawson. 1996. "Coat Color Genetics of Peromyscus: V. California Blonde, a New Recessive Mutation in the Deer Mouse." *The Journal of Heredity* 87 (5): 403–6.

Roth, V Louise, and Maryrose S Klein. 1986. "Maternal Effects of Body Size of Large Insular Peromyscus Maniculatus : Evidence from Embryo Transfer Experiments." *Journal of Mammalogy* 67 (1): 37–45.

Schmidt-Nielsen, K. 1984. *Scaling: Why Is Animal Size so Important?* Cambridge University Press.

Sekhon, Jasjeet S. 2011. "Multivariate and Propensity Score Matching Software with Automated Balance Optimization: The Matching Package for R." *Journal of Statistical Software* 42: 1–52.

Sidlauskas, Brian, Ganeshkumar Ganapathy, Einat Hazkani-Covo, Kristin P Jenkins, Hilmar Lapp, Lauren W McCall, Samantha Price, Ryan Scherle, Paula a Spaeth, and David M Kidd. 2010. "Linking Big: The Continuing Promise of Evolutionary Synthesis." *Evolution; International Journal of Organic Evolution* 64 (4): 871–80. doi:10.1111/j.1558-5646.2009.00892.x.

Smith, Felisa A., S. Kathleen Lyons, S. K. Morgan Ernest, Kate E. Jones, Dawn M. Kaufman, Tamar Dayan, Pablo a. Marquet, James H. Brown, and John P. Haskell. 2003. "Body Mass of Late Quaternary Mammals." *Ecology* 84 (12): 3403–3403. doi:10.1890/02-9003.

Song, Sen, Liang Liu, Scott V Edwards, and Shaoyuan Wu. 2012. "Resolving Conflict in Eutherian Mammal Phylogeny Using Phylogenomics and the Multispecies Coalescent Model." *Proceedings of the National Academy of Sciences of the United States of America* 109 (37): 14942–47. doi:10.1073/pnas.1211733109.

Steppan, Scott, Ronald Adkins, and Joel Anderson. 2004. "Phylogeny and Divergence-Date Estimates of Rapid Radiations in Muroid Rodents Based on Multiple Nuclear Genes." *Systematic Biology* 53 (4). Oxford University Press: 533–53.

Swofford, David L. 2002. "Phylogenetic Analysis Using Parsimony * (and Other Methods). Version 4." *Options*. Sinauer Associates, Sunderland, Massachusetts. doi:10.1159/000170955.

Taylor, Zachary S, and Susan M G Hoffman. 2010. "Mitochondrial DNA Genetic Structure Transcends Natural Boundaries in Great Lakes Populations of Woodland Deer Mice (*Peromyscus Maniculatus Gracilis*)." *Canadian Journal of Zoology* 88 (4): 404–15. doi:10.1139/Z10-010.

———. 2011. "Microsatellite Genetic Structure and Cytonuclear Discordance in Naturally Fragmented Populations of Deer Mice (Peromyscus Maniculatus)." *The Journal of Heredity* 103 (1): 71–79. doi:10.1093/jhered/esr100.

Therneau, Terry M, Beth Atkinson, and Brian D Ripley. 2006. "Rpart: Recursive Partitioning." *Rpart Package Manual*.

Therneau, Terry M, and Elizabeth J Atkinson. 1997. "An Introduction to Recursive Partitioning Using the Rpart Routine." *Stats* 116: 1–52.

Uthicke, S, and J A H Benzie. 2003. "Gene Flow and Population History in High Dispersal Marine Invertebrates: Mitochondrial DNA Analysis …." *Molecular Ecology* 12: 2636–48. doi:10.1046/j.1365-294X.2003.01954.x.

Van Valen, L. 1973. "Pattern and the Balance of Nature." *Evolutionary Theory* 1 (July): 31–49.

Walker, PL. 1980. "Archaeological Evidence for the Recent Extinction of Three Terrestrial Mammals on San Miguel Island." In *The Proceedings of a Multidisciplinary Symposium, DM Power (ed.). Santa Barbara Natural History Museum, Santa Barbara, California*, 703–17.

Wallace, Alfred Russel. 1876. "The Geographical Distribution of Animals with a Study of the Relations of Living and Extinct Faunas as Elucidating the Past Changes of the Earth's Surface". London:[sn].

Wang, Yanping, Shuihua Chen, and Ping Ding. 2011. "Testing Multiple Assembly Rule Models in Avian Communities on Islands of an Inundated Lake, Zhejiang Province, China." *Journal of Biogeography* 38 (7): 1330–44. doi:10.1111/j.1365-2699.2011.02502.x.

Wayne, Robert K, Sarah B George, Dennis Gilbert, Paul W Collins, Steven D Kovach, Derek Girman, and Niles Lehman. 1991. "A Morphologic and Genetic Study of the Island Fox, Urocyon Littoralis." *Evolution* 45 (8). JSTOR: 1849–68.

Weber, Jesse N, Maureen B Peters, Olga V Tsyusko, Catherine R Linnen, Cris Hagen, Nancy a Schable, Tracey D Tuberville, et al. 2010. "Five Hundred Microsatellite Loci for Peromyscus." *Conservation Genetics (Print)* 11 (3): 1243–46. doi:10.1007/s10592-009-9941-x.

Whittaker, RJ, and JM Fernández-Palacios. 2007. *Island Biogeography: Ecology, Evolution, and Conservation*.

Wilson, Amy. 2008. "The Role of Insularity in Promoting Intraspecific Differentiation in Song Sparrows". University of British Columbia.

Wilson, D.E., and D.A.M. Reeder. 2005. *Mammal Species of the World: A Taxonomic and Geographic Reference*. Johns Hopkins University Press.

Yang, Dou-Shuan, and G J Kenagy. 2009. "Nuclear and Mitochondrial DNA Reveal Contrasting Evolutionary Processes in Populations of Deer Mice (Peromyscus Maniculatus)." *Molecular Ecology* 18 (24): 5115–25. doi:10.1111/j.1365-294X.2009.04399.x.

Biography

Paul Durst was born in Hayward, California in 1983. As an undergraduate, he attended the University of California, Berkeley, graduating with a B.A. in Integrative Biology in 2006. After two years of working as a field technician and high school instructor, he came to Duke University. While at Duke, he has received several honors and fellowships including the James B. Duke Fellowship (2008), Honorable Mention in the NSF Graduate Research Fellowship program (2009), a NESCent Graduate Student Fellowship (2010), an NSF Doctoral Dissertation Improvement Grant (2013) and most recently, the SPIRE Postdoctoral Fellowship at the University of North Carolina at Chapel Hill, where he will begin work in the fall.

Peer-reviewed publications:

McClain, C.R., P.A.P. Durst, A.G. Boyer, C.D. Francis, 2013. Unravelling the determinants of insular body size shifts. *Biology Letters, 9(1),* doi:10.1098/rsbl.2012.0989

Durst, P.A.P. & V.L. Roth, 2012. Classification-tree methods provide a multifactorial approach to predicting insular body-size evolution in rodents. *The American Naturalist, 179(4)*, 545-553, doi: 10.1086/664611